

Murine Bin-1 mutant
 Murine Bin-1 mutant
 Murine Bin-1 mutant
 Novel human diagno
 Novel human diagno
 Murine Bcl-2 inter
 Bcl2 polypeptide B
 Mammalian Bim Bcl-
 BIM BH3 consensus
 Human polypeptide
 Human polypeptide
 Human polypeptide
 Human dunce-like p
 pPDE46 human dunce
 Novel human diagno
 Novel human diagno
 Amino acid sequenc
 Amino acid sequenc
 Amino acid sequenc
 Amino acid sequenc
 Amino acid sequenc
 Novel human diagno
 Human phosphodiester
 HSV-2 strain SB5 C
 HSV-2 strain SB5 C
 HSV-2 strain SB5 C
 Drosophila melanog
 WASP homolog prote
 Amino acid sequenc
 Human N-methyl-D-a
 Human N-methyl-D-a
 Human N-methyl-D-a
 Novel human diagno
 Human ataxin-2 lik
 Human polypeptide,
 Human protein sequ
 Novel human diagno
 Human protein SEO

Run on: August 16, 2002, 09:56:37 ; Search time 30.56 Seconds
(without alignments)
719.654 Million cell updates/sec

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Database :

A_Geneseq_032802.*

1:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Length				
1	1065	100.0	198	20	AAW98158	Human Bcl-2 intera	
2	908	85.3	196	20	AAW98156	Murine Bcl-2 inter	
3	696	65.4	138	20	AAW98164	Human Bim-L mutant	
4	696	65.4	138	20	AAW98157	Human Bcl-2 intera	
5	691	64.9	138	20	AAW98105	Human Bim-L mutant	
6	691	64.9	138	20	AAW98166	Human Bim-L mutant	
7	691	64.9	138	20	AAW98168	Human Bim-L mutant	
8	690	64.8	138	20	AAW98167	Human Bim-L mutant	
9	589	55.3	140	20	AAW98135	Murine Bcl-2 inter	
10	584	54.8	140	20	AAW98160	Murine Bim-L mutan	
11	584	54.8	140	20	AAW98161	Murine Bim-L mutan	

RESULT 1

AA
AC AAW98158:

Human Bcl-2 interacting mediator of cell death Bim-EL isoform.

OS Homo sapiens.

FT /note= "dynein light chain binding region"

```

11 region
12 FT
13 /note= "hydrophobic region"
14 121.1455

```

PD 25-MAR-1999.

PR 24-SEP-1997; 97AU-0009373.

XX
XX

CC treatment or prophylaxis in conditions such as cancer and deletion
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
 CC expression of Bim activity is useful in regulating inhibition or
 CC prevention of cell death or degeneration such as under cytotoxic
 CC conditions during e.g. gamma-irradiation and chemotherapy or during
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
 CC hypoxia, degenerative diseases or for prolonging the survival of
 CC cells being transplanted for treatment of disease. Since Bim is
 CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 XX
 SQ Sequence 196 AA;

Query Match 85.3%; Score 908; DB 20; Length 196;
 Best Local Similarity 86.5%; Pred. No. 2e-78; 11; Indels 6; Gaps 2;
 Matches 173; Conservative 10; Mismatches 11; Indels 6; Gaps 2;
 QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTEPQGNHGEGDSCPHGSP 60
 DB 1 makqpsdvssecdregglqpaerppqlrpgaptslqtepgnqd----gedrcphgsp 56
 QY 61 QGPLAPPASGPFATRSPLIFIMRRSSLLSRSSSGYFSDTRSPAPMSCDKSTQTPSP 120
 DB 57 qgplappasgpfatrsplifivrssllsrsssgyfsfdrspapmscdkstqtpsp 116
 QY 121 COAFNHLYLSAMASMRQA--EPADMPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDH 178
 DB 117 cqafnhylsamasirgsgeepdrpeiriaqelrrigdefnetyrrrvfandyreaedh 176
 QY 179 PRMVLRLRYIVRLVWRMH 198
 DB 177 pqmvlrlrlfvlrvwrh 196

RESULT 3
 AAW98164
 ID AAW98164 standard; Protein; 138 AA.
 AC AAW98164;
 XX
 DT 05-JUL-1999 (first entry)
 DE Human Bim-L mutant D51G.
 KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; human; cancer; autoimmune disease; therapy;
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;
 KW cytotoxicity; agonist.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Binding-site 42..71
 FT /note= "dynein light chain binding region, present
 FT only in Bim-L, Bim-EL"
 FT
 FT Misc-difference 51
 FT /note= "Gly replaces Asp of wild-type Bim-L"
 FT Region 94..102
 FT /note= "BH3"
 FT Region 123..137
 FT /note= "hydrophobic region"
 XX
 PN WO9914321-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 17-SEP-1998; 98WO-AU00772.
 XX
 XX 24-SEP-1997; 97AU-0009373.
 PR 17-SEP-1997; 97AU-0009263.

XX
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX
 DR WPI; 1999-244030/20.
 XX
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment
 XX
 PS Claim 25; Page -; 145pp; English.
 XX
 CC The present sequence is a D51G mutant of the long form (L) of
 CC human Bim, or Bcl-2 interacting mediator of cell death (see also
 CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is
 CC capable of inducing cell death (apoptosis) and acts as a
 CC 'death-ligand' for certain members of the pro-survival Bcl-2
 CC family. Binding the dynein light chain regulates the pro-apoptotic
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
 CC not bind to dynein light chain, is a much more potent killer than
 CC either Bim-L or Bim-EL. The invention provides variants (see
 CC AAW98159-58) of murine and human Bim-L and Bim-EL that cannot bind,
 CC couple or otherwise associate with a dynein light chain owing to
 CC amino acid addition, substitution and/or deletion within the region
 CC which binds the dynein light chain. The variants act as agonists
 CC of Bim function, interacting with anti-apoptotic molecules such as
 CC Bcl-2 to prevent their functional activity, thereby promoting
 CC apoptosis, and can be used e.g. in the treatment of cancer or to
 CC deplete autoreactive lymphocytes in autoimmune diseases. Since Bim
 CC is expressed in germ cells, modulating its expression or activity
 CC may be useful as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 CC Note. The present sequence is not shown in the specification but
 CC is derived from the human Bim-L sequence given on page 100-101.
 XX
 SQ Sequence 138 AA;

Query Match 65.4%; Score 696; DB 20; Length 138;
 Best Local Similarity 69.7%; Pred. No. 1.8e-58;
 Matches 138; Conservative 0; Mismatches 0; Indels 60; Gaps 1;
 QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTEPQGNHGEGDSCPHGSP 60
 DB 1 makqpsdvssecdregglqpaerppqlrpgaptslqtepgnqd----gedrcphgsp 56
 QY 61 QGPLAPPASGPFATRSPLIFIMRRSSLLSRSSSGYFSDTRSPAPMSCDKSTQTPSP 120
 DB 42 -----drspapmscdkstqtpsp 60
 QY 121 COAFNHLYLSAMASMRQAEPADMPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDH 180
 DB 61 cqafnhylsamasirgsgeepdrpeiriaqelrrigdefnetyrrrvfandyreaedh 176
 QY 181 MVILRLRYIVRLVWRMH 198
 DB 121 mvilrlryivrlvwrh 138
 RESULT 4
 AAW98157
 ID AAW98157 standard; Protein; 138 AA.
 XX
 AC AAW98157;
 XX
 DT 05-JUL-1999 (first entry)
 DE Human Bcl-2 interacting mediator of cell death Bim-L isoform.
 XX
 KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; human; cancer; autoimmune disease;
 KW degenerative disease; therapy; contraceptive; splice variant;

KW isoform.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Binding-site 42..71 /note= "dynein light chain binding region, present
 FT only in Bim-L, Bim-EL"
 FT Region 92..100 /note= "BH3"
 FT Region. 121..135 /note= "hydrophobic region"
 FT
 FT
 XX WO9914321-A1.
 PN
 XX
 XX 25-MAR-1999.
 XX
 XX 17-SEP-1998; 98WO-AU00772.
 XX
 XX 24-SEP-1997; 97AU-0009373.
 PR 17-SEP-1997; 97AU-0009263.
 XX
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PA
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX WPI: 1999-244030/20.
 XX N-PSDB; AAX24996.
 DR
 DR
 XX
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment
 XX
 XX Claim 15; Page 100-101; 145pp; English.
 PS
 XX
 CC The present sequence is the long form (L) of human Bim, or Bcl-2
 CC interacting mediator of cell death, a novel member of the Bcl-2
 CC family that is capable of inducing cell death (apoptosis) and which
 CC acts as a 'death-ligand' for certain members of the pro-survival
 CC Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology
 CC region which it encompasses is BH3. It is the only BH3-only protein
 CC for which splice variants exist. These result in the expression of
 CC a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL. cDNAs (see
 CC AAX24996-97) encoding human Bim-L and Bim-EL (see AAW98165) were
 CC obtained from embryo and liver cDNA libraries using mouse Bim cDNA
 CC as probe. Murine Bim-S, Bim-L and Bim-EL isoforms are also
 CC provided (see AAW98154-56). Binding of the dynein light chain was shown
 CC to regulate the pro-apoptotic activity of Bim. Bim-S, the splice
 CC variant which does not bind to dynein light chain, is a much more
 CC potent killer than either Bim-L or Bim-EL. The invention provides
 CC variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that
 CC cannot bind, couple or otherwise associate with a dynein light
 CC chain. The identification of Bim permits the identification and
 CC rational design of a range of products for use in therapy,
 CC diagnosis, antibody generation and involving modulation of
 CC physiological cell death. These therapeutic molecules may act
 CC as either antagonists or agonists of Bim's function and will be
 CC useful in cancer, autoimmune or degenerative disease therapy.
 CC Increased Bim expression or Bim activity is useful, e.g. for
 CC treatment or prophylaxis in conditions such as cancer and deletion
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
 CC expression of Bim activity is useful in regulating inhibition or
 CC prevention of cell death or degeneration such as under cytotoxic
 CC conditions during e.g. gamma-irradiation and chemotherapy or during
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
 CC hypoxia, degenerative diseases or for prolonging the survival of
 CC cells being transplanted for treatment of disease. Since Bim is
 CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 XX Sequence 138 AA;

Query Match 65.4%; Score 696; DB 20; Length 138;
 Best Local Similarity 69.7%; Pred. No. 1.8e-58;
 Matches 138; Conservative 0; Mismatches 0; Indels 60; Gaps 1;
 QY 1 MAKQPSVSSCEDREGLOLQPAERPPQLRPGAPTSLOTPEOGNPEGNHGGEDSCPHGSP 60
 DB 1 makqpsdvsscedregqlgpaerppqlrpgaptslqtpeq----- 41
 QY 61 QGPLAPPASPCPPFATRSPLFIFMRSSILLSRSSSGYFSFDTDRSPAPMSCDKSTQTSPP 120
 DB 42 -----drspapmscdkstqtspp 60
 QY 121 CQAFNHVLSAMASMRQAEPADMRPEIWTQBELRRIGDEFNAYYARRVFLNNYQAAEDHPR 180
 DB 61 cqafnhvlsamasmrqaeapadmripeiwaqlerriqdefnayyarrvflnnnyaedhpr 120
 QY 181 MVILRLRYIVRLVWRMH 198
 DB 121 mvilrllryivrlvwrmh 138
 RESULT 5
 AAW98165
 ID AAW98165 standard; Protein; 138 AA.
 XX
 AC AAW98165;
 XX
 DT 05-JUL-1999 (first entry)
 XX
 DE Human Bim-L mutant S53P.
 XX
 KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; human; cancer; autoimmune disease; therapy;
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;
 KW cytotoxicity; agonist.
 XX
 OS Homo sapiens.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Binding-site 42..71 /note= "dynein light chain binding region, present
 FT only in Bim-L, Bim-EL"
 FT Misc-difference 53 /note= "Pro replaces Ser of wild-type Bim-L"
 FT Region 94..102 /note= "BH3"
 FT Region 123..137 /note= "hydrophobic region"
 FT
 FT WO9914321-A1.
 XX
 XX 25-MAR-1999.
 XX
 XX 17-SEP-1998; 98WO-AU00772.
 XX
 XX 24-SEP-1997; 97AU-0009373.
 PR 17-SEP-1997; 97AU-0009263.
 XX
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PA
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX WPI: 1999-244030/20.
 DR
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment
 XX
 XX Claim 25; Page -; 145pp; English.
 XX
 CC The present sequence is a S53P mutant of the long form (L) of

CC human Bim, or Bcl-2 interacting mediator of cell death (see also
 CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is
 CC capable of inducing cell death (apoptosis) and acts as a
 CC 'death-ligand' for certain members of the pro-survival Bcl-2
 CC family. Binding the dynein light chain regulates the pro-apoptotic
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
 CC not bind to dynein light chain, is a much more potent killer than
 CC either Bim-L or Bim-EL. The invention provides variants (see
 CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
 CC couple or otherwise associate with a dynein light chain owing to
 CC amino acid addition, substitution and/or deletion within the region
 CC which binds the dynein light chain. The variants act as agonists
 CC of Bim function, interacting with anti-apoptotic molecules such as
 CC Bcl-2 to prevent their functional activity, thereby promoting
 CC apoptosis, and can be used e.g. in the treatment of cancer or to
 CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
 CC is expressed in germ cells, modulating its expression or activity
 CC may be useful as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 CC Note. The present sequence is not shown in the specification but
 CC is derived from the human Bim-L sequence given on page 100-101.
 XX
 XX Sequence 138 AA;
 SQ

Query Match 64.9%; Score 691; DB 20; Length 138;
 Best Local Similarity 69.2%; Pred. No. 5.4e-58;
 Matches 137; Conservative 0; Mismatches 1; Indels 60; Gaps 1;

QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTEPGNHNHGGSDSCPHGSP 60
 DB 1 makqpsdvssecdregrlqpaerppqlrpgaptslqtpeq----- 41
 QY 61 QGPLAPPASGPPATRSPLFIEMRRSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120
 DB 42 -----drspapmscdktpqtsp 60
 QY 121 COAFNHYLSAMASMRQAEPADMRPEIWIQELRRIGDEFNAYYARRVFLNNYQAEDHPR 180
 DB 61 cqafnhylsamasmrqaeapadmrpeiwiqelrrigdefnayyarrvflnnyqaedhpr 120
 QY 181 MVILRLRLRYIVRLVWRMH 198
 DB 121 mvilrlrlryivrlvwrmh 138

RESULT 6
 AAW98166 standard; Protein; 138 AA.
 AC AAW98166;
 XX
 XX 05-JUL-1999 (first entry)
 XX Human Bim-L mutant T54A.
 DE
 XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; human; cancer; autoimmune disease; therapy;
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;
 KW cytotoxicity; agonist.
 XX
 XX Homo sapiens.
 OS Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Binding-site 42..71
 FT /note= "dynein light chain binding region, present
 FT only in Bim-L, Bim-EL"
 FT Misc-difference 54
 FT /note= "Ala replaces Thr of wild-type Bim-L"
 FT Region 94..102
 FT /note= "BH3"
 FT Region 123..137

FT /note= "hydrophobic region"
 XX
 PN W09914321-A1.
 XX
 PD 25-MAR-1999.
 XX
 XX 17-SEP-1998; 98WO-AU00772.
 PF
 XX 24-SEP-1997; 97AU-0009373.
 PR
 XX 17-SEP-1997; 97AU-0009263.
 XX
 XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 PA
 XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX
 XX WPI; 1999-244030/20.
 DR
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment
 PT
 XX Claim 25; Page -; 145pp; English.
 XX
 CC The present sequence is a T54A mutant of the long form (L) of
 CC human Bim, or Bcl-2 interacting mediator of cell death (see also
 CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is
 CC capable of inducing cell death (apoptosis) and acts as a
 CC 'death-ligand' for certain members of the pro-survival Bcl-2
 CC family. Binding the dynein light chain regulates the pro-apoptotic
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
 CC not bind to dynein light chain, is a much more potent killer than
 CC either Bim-L or Bim-EL. The invention provides variants (see
 CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
 CC couple or otherwise associate with a dynein light chain owing to
 CC amino acid addition, substitution and/or deletion within the region
 CC which binds the dynein light chain. The variants act as agonists
 CC of Bim function, interacting with anti-apoptotic molecules such as
 CC Bcl-2 to prevent their functional activity, thereby promoting
 CC apoptosis, and can be used e.g. in the treatment of cancer or to
 CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
 CC is expressed in germ cells, modulating its expression or activity
 CC may be useful as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 CC Note. The present sequence is not shown in the specification but
 CC is derived from the human Bim-L sequence given on page 100-101.
 XX
 XX Sequence 138 AA;
 SQ

Query Match 64.9%; Score 691; DB 20; Length 138;
 Best Local Similarity 69.2%; Pred. No. 5.4e-58;
 Matches 137; Conservative 0; Mismatches 1; Indels 60; Gaps 1;

QY 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLSQTEPGNHNHGGSDSCPHGSP 60
 DB 1 makqpsdvssecdregrlqpaerppqlrpgaptslqtpeq----- 41
 QY 61 QGPLAPPASGPPATRSPLFIEMRRSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120
 DB 42 -----drspapmscdkstaqtsp 60
 QY 121 COAFNHYLSAMASMRQAEPADMRPEIWIQELRRIGDEFNAYYARRVFLNNYQAEDHPR 180
 DB 61 cqafnhylsamasmrqaeapadmrpeiwiqelrrigdefnayyarrvflnnyqaedhpr 120
 QY 181 MVILRLRLRYIVRLVWRMH 198
 DB 121 mvilrlrlryivrlvwrmh 138

RESULT 7
 AAW98168 standard; Protein; 138 AA.
 ID AAW98168

XX AC AAW98168;
 XX XX
 XX 05-JUL-1999 (first entry)
 XX DE Human Bim-L mutant N65S.
 XX XX
 XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; human; cancer; autoimmune disease; therapy;
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;
 KW cytotoxicity; agonist.
 XX XX
 OS Homo sapiens.
 OS Synthetic.
 XX XX
 FH Key Location/Qualifiers
 FT Binding-site 42..71
 FT /note= "dynein light chain binding region, present
 FT only in Bim-L, Bim-EL"
 FT
 FT Misc-difference 65
 FT /note= "Ser replaces Asn of wild-type Bim-L"
 FT 94..102
 FT /note= "BH3"
 FT 123..137
 FT /note= "hydrophobic region"
 FT
 XX WO9914321-A1.
 XX XX
 XX 25-MAR-1999.
 XX XX
 XX 17-SEP-1998; 98WO-AU000772.
 XX XX
 XX 24-SEP-1997; 97AU-0009373.
 PR 17-SEP-1997; 97AU-0009263.
 XX XX
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX XX
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX XX
 DR WPI; 1999-244030/20.
 XX XX
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 FT treatment
 FT
 PS Claim 25; Page -; 145pp; English.
 XX XX
 CC The present sequence is a N65S mutant of the long form (L) of
 CC human Bim, or Bcl-2 interacting mediator of cell death (see also
 CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is
 CC capable of inducing cell death (apoptosis) and acts as a
 CC 'death-ligand' for certain members of the pro-survival Bcl-2
 CC family. Binding the dynein light chain regulates the pro-apoptotic
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
 CC not bind to dynein light chain, is a much more potent killer than
 CC either Bim-L or Bim-EL. The invention provides variants (see
 CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
 CC couple or otherwise associate with a dynein light chain owing to
 CC amino acid addition, substitution and/or deletion within the region
 CC which binds the dynein light chain. The variants act as agonists
 CC of Bim function, interacting with anti-apoptotic molecules such as
 CC Bcl-2 to prevent their functional activity, thereby promoting
 CC apoptosis, and can be used e.g. in the treatment of cancer or to
 CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
 CC is expressed in germ cells, modulating its expression or activity
 CC may be useful as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 CC Note. The present sequence is not shown in the specification but
 CC is derived from the human Bim-L sequence given on page 100-101.
 XX XX
 SQ Sequence 138 AA;

Query Match 64.9%; Score 691; DB 20; Length 138;
 Best Local Similarity 69.2%; Pred. No. 5.4e-58;
 Matches 137; Conservative 1; Mismatches 0; Indels 60; Gaps 1;
 QY 1 MAKQPSDVSSCEDREGROLOPAERPPQLRPGAPTSLQTEPOGNEGNHGGEDSCPHGSP 60
 DB 1 makqpsdvsscedregrolopaerppqlrpgaptslqtepq----- 41
 QY 61 QGPLAPPASPGPFATRSPLFIEMRRSSLLSRSSSGYFSDTRDSPAPMSCDKSTQTPSP 120
 DB 42 -----drspapmscdkstqtpsp 60
 QY 121 COAFNHYLSAMASMRQAEPPADMRPEIWTAEQLRIGDGFNAYARRVFLNNYQAAEDHPR 180
 DB 61 cqafshylsamasmrqaepadmpeiqaelrrigdfefnayarrvflnnynqaaedhpr 120
 QY 181 MVILRLRYIVRLVWRMH 198
 DB 121 mvilrlryivrlvwrmh 138
 RESULT 8
 AAW98167
 ID AAW98167 standard; Protein; 138 AA.
 XX AC AAW98167;
 XX XX
 XX 05-JUL-1999 (first entry)
 XX XX
 XX Human Bim-L mutant T54I.
 KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
 KW cell cycle; human; cancer; autoimmune disease; therapy;
 KW contraceptive; splice variant; isoform; mutant; dynein light chain;
 KW cytotoxicity; agonist.
 XX XX
 OS Homo sapiens.
 OS Synthetic.
 XX XX
 FH Key Location/Qualifiers
 FT Binding-site 42..71
 FT /note= "dynein light chain binding region, present
 FT only in Bim-L, Bim-EL"
 FT
 FT Misc-difference 54
 FT /note= "Ile replaces Thr of wild-type Bim-L"
 FT 94..102
 FT /note= "BH3"
 FT 123..137
 FT /note= "hydrophobic region"
 FT
 FT WO9914321-A1.
 XX XX
 XX 25-MAR-1999.
 XX XX
 XX 17-SEP-1998; 98WO-AU000772.
 XX XX
 XX 24-SEP-1997; 97AU-0009373.
 PR 17-SEP-1997; 97AU-0009263.
 XX XX
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX XX
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX XX
 DR WPI; 1999-244030/20.
 XX XX
 XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 FT treatment
 FT
 PS Claim 25; Page -; 145pp; English.
 XX XX
 CC The present sequence is a T54I mutant of the long form (L) of
 CC human Bim, or Bcl-2 interacting mediator of cell death (see also

CC AAW98157), a novel member of the Bcl-2 family. Native Bim-L is
 CC capable of inducing cell death (apoptosis) and acts as a
 CC 'death-ligand' for certain members of the pro-survival Bcl-2
 CC family. Binding the dynein light chain regulates the pro-apoptotic
 CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
 CC not bind to dynein light chain, is a much more potent killer than
 CC either Bim-L or Bim-EL. The invention provides variants (see
 CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
 CC couple or otherwise associate with a dynein light chain owing to
 CC amino acid addition, substitution and/or deletion within the region
 CC which binds the dynein light chain. The variants act as agonists
 CC of Bim function, interacting with anti-apoptotic molecules such as
 CC Bcl-2 to prevent their functional activity, thereby promoting
 CC apoptosis, and can be used e.g. in the treatment of cancer or to
 CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
 CC is expressed in germ cells, modulating its expression or activity
 CC may be useful as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 CC Note. The present sequence is not shown in the specification but
 CC is derived from the human Bim-L sequence given on page 100-101.
 XX
 SQ Sequence 138 AA;

Query Match 64.8%; Score 690; DB 20; Length 138;
 Best Local Similarity 69.2%; Pred. No. 6.7e-58;
 Matches 137; Conservative 0; Mismatches 1; Indels 60; Gaps 1;

QY 1 MAKQSDVSSCDREGRLQPAERPPQLRPGAPTSLSQTEPQGNPEGHGEGDSCPHGSP 60
 DB 1 makqsdvssecdregrlqpaerppqlrpgaptslqtepq----- 41
 QY 61 QGFLAPPASGPFATRSPLFIEMRRSLLSRSSGYSFDTDRSPAPMSCDKSTQTPSP 120
 DB 42 -----drspapmscdksiqtpsp 60

QY 121 COAFNHLSAMASMRQAEADMRPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDHPR 180
 DB 61 cqafnhlsamsmrqaepadmrpeilwiqaqlrrigdefnayyarrvflnnyyqaedhpr 120

QY 181 MWILRLRYIVRLVWRMH 198
 DB 121 mwilrlryivrlvwrmh 138

RESULT 9
 ID AAW98155 standard; Protein: 140 AA.
 AC AAW98155;
 DT 05-JUL-1999 (first entry)
 DE Murine Bcl-2 interacting mediator of cell death Bim-L isoform.
 XX Bim-L; Bcl-2 interacting mediator of cell death; Bim-L isoform.
 KW cell cycle; mouse; cancer; autoimmune disease;
 KW degenerative disease; therapy; contraceptive; splice variant;
 KW isoform.
 XX Mus musculus.
 XX Key Location/Qualifiers
 FH Binding-site 42..71
 FT /note= "dynein light chain binding region, present
 FT only in Bim-L, Bim-EL"
 FT Region 94..102
 FT /note= "BH3"
 FT Region 123..137
 FT /note= "hydrophobic region"
 XX WO9914321-A1.
 PN
 XX

PD 25-MAR-1999.
 XX
 PF 17-SEP-1998; 98WO-AU00772.
 XX
 PR 24-SEP-1997; 97AU-0009373.
 PR 17-SEP-1997; 97AU-0009263.
 XX
 PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
 XX
 PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
 PI Puthalakath H, Strasser A;
 XX
 DR WPI; 1999-244030/20.
 DR N-PSDB; AAX24994.
 XX
 PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
 PT treatment
 XX
 PS Claim 11; Page 95-96; 145pp; English.
 XX
 CC The present sequence is the long form (L) of murine Bim, or Bcl-2
 CC interacting mediator of cell death, a novel member of the Bcl-2
 CC family that is capable of inducing cell death (apoptosis) and which
 CC acts as a 'death-ligand' for certain members of the pro-survival
 CC Bcl-2 family. Bim is a BH3-only protein, as the only Bcl-2 homology
 CC region which it encompasses is BH3. It is the only BH3-only protein
 CC for which splice variants exist. These result in the expression of
 CC a variety of isoforms, i.e. Bim-S, Bim-L and Bim-EL (see AAW98154-56).
 CC cDNAs encoding these murine Bim isoforms (see AAX24993-95) were
 CC obtained from a T lymphoma cDNA library using human recombinant
 CC Bcl-2 protein. Human Bim-L and Bim-EL isoforms have also been
 CC identified (see AAW98157-58). Binding the dynein light chain was shown
 CC to regulate the pro-apoptotic activity of Bim. Bim-S, the splice
 CC variant which does not bind to dynein light chain, is a much more
 CC potent killer than either Bim-L or Bim-EL. The invention provides
 CC variants (see AAW98159-68) of murine and human Bim-L or Bim-EL that
 CC cannot bind, couple or otherwise associate with a dynein light
 CC chain. The identification of Bim permits the identification and
 CC rational design of a range of products for use in therapy.
 CC diagnosis, antibody generation and involving modulation of
 CC physiological cell death. These therapeutic molecules may act
 CC as either antagonists or agonists of Bim's function and will be
 CC useful in cancer, autoimmune or degenerative disease therapy.
 CC Increased Bim expression or Bim activity is useful, e.g. for
 CC treatment or prophylaxis in conditions such as cancer and deletion
 CC of autoreactive lymphocytes in autoimmune disease. Decreased Bim
 CC expression of Bim activity is useful in regulating inhibition or
 CC prevention of cell death or degeneration such as under cytotoxic
 CC conditions during e.g. gamma-irradiation and chemotherapy or during
 CC HIV/AIDS or other viral infections, ischemia, myocardial infarction,
 CC hypoxia, degenerative diseases or for prolonging the survival of
 CC cells being transplanted for treatment of disease. Since Bim is
 CC expressed in germ cells, modulating Bim expression or Bim activity
 CC is useful, e.g. as a contraceptive or method of sterilization by
 CC preventing generation of fertile sperm.
 XX
 SQ Sequence 140 AA;

Query Match 55.3%; Score 589; DB 20; Length 140;
 Best Local Similarity 60.0%; Pred. No. 2.7e-48;
 Matches 120; Conservative 8; Mismatches 10; Indels 62; Gaps 2;

QY 1 MAKQSDVSSCDREGRLQPAERPPQLRPGAPTSLSQTEPQGNPEGHGEGDSCPHGSP 60
 DB 1 makqsdvssecdregrlqpaerppqlrpgaptslqtepq----- 41
 QY 61 QGFLAPPASGPFATRSPLFIEMRRSLLSRSSGYSFDTDRSPAPMSCDKSTQTPSP 120
 DB 42 -----drspapmscdksiqtpsp 60
 QY 121 COAFNHLSAMASMRQAEADMRPEIWIQAELRRIGDEFNAYYARRVFLNNYQAEDH 178
 XX

Db 61 cqafnhylsamarsiqseepdlrpeiriaeqlrrigdefnetytrrvfandyreaedh 120

Qy 179 PRMVLRLRLRYIVRLVWRMH 198
l:||||:||||:| ||||| |

Db 121 pqmvilqlrlfrlrvwrh 140

RESULT 10
AAW98160
ID AAW98160 standard; Protein; 140 AA.
XX AC AAW98160;
XX DT 05-JUL-1999 (first entry)
XX DE Murine Bim-L mutant S53P.
XX KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW cell cycle; mouse; cancer; autoimmune disease; therapy;
KW contraceptive; splice variant; isoform; mutant; dynein light chain;
KW cytotoxicity; agonist.
XX OS Mus musculus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX Binding-site 42..71
/note= "dynein light chain binding region, present
only in Bim-L, Bim-EL"
XX FT Misc-difference 53
/note= "Pro replaces Ser of wild-type Bim-L"
XX FT Region 94..102
/note= "BH3"
XX FT Region 123..137
/note= "hydrophobic region"
XX FT
XX PN W09914321-A1.
XX PD 25-MAR-1999.
XX PF 17-SEP-1998; 98WO-AU00772.
XX PR 24-SEP-1997; 97AU-0009373.
XX PR 17-SEP-1997; 97AU-0009263.
XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI Puthalakath H, Strasser A;
XX DR WPI; 1999-244030/20.
XX PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT treatment
XX PS Claim 25; Page -; 145pp; English.
XX CC The present sequence is a S53p mutant of the long form (L) of
CC murine Bim, or Bcl-2 interacting mediator of cell death (see also
CC AAW98155), a novel member of the Bcl-2 family. Native Bim-L is
CC capable of inducing cell death (apoptosis) and acts as a
CC 'death-ligand' for certain members of the pro-survival Bcl-2
CC family. Binding the dynein light chain regulates the pro-apoptotic
CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
CC not bind to dynein light chain, is a much more potent killer than
CC either Bim-L or Bim-EL. The invention provides variants (see
CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
CC couple or otherwise associate with a dynein light chain owing to
CC amino acid addition, substitution and/or deletion within the region
CC which binds the dynein light chain. The variants act as agonists
CC of Bim function, interacting with anti-apoptotic molecules such as
CC Bcl-2 to prevent their functional activity, thereby promoting
CC apoptosis, and can be used e.g. in the treatment of cancer or to

CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
CC is expressed in germ cells, modulating its expression or activity
CC may be useful as a contraceptive or method of sterilization by
CC preventing generation of fertile sperm.
CC Note. The present sequence is not shown in the specification but
CC is derived from the murine Bim-L sequence given on page 95-96.
XX SQ Sequence 140 AA;
Query Match 54.8%; Score 584; DB 20; Length 140;
Best Local Similarity 59.5%; Pred. No. 8.1e-48;
Matches 119; Conservative 8; Mismatches 11; Indels 62; Gaps 2;
Qy 1 MAKQPSDVSECDREGRLQPAERPPQLRPGAPTSLOTEPGQGNPEGNHGGSDSCPHGSP 60
|||||:|||||:| ||||| |
Db 1 makpsdvssecdreggqlqpaerppqlrpgaptsltepq----- 41
Qy 61 QGPIAPPASPGPPFATRSPLFIEMRRSSLLSRSSGYSFDTDRSPAPMSCDKSTQTPSP 120
|||||:|||||:| ||||| |
Db 42 -----drspapmscdkptqtpsp 60
Qy 121 CQAFNHVLSAMSRQA--EPADMRPEITWIAQELRLRIGDENAYYARRVFLNYYQAAEDH 178
|||||:|||||:| ||||| |
Db 61 cqafnhylsamarsiqseepdlrpeiriaeqlrrigdefnetytrrvfandyreaedh 120
Qy 179 PRMVLRLRLRYIVRLVWRMH 198
l:||||:||||:| ||||| |
Db 121 pqmvilqlrlfrlrvwrh 140
RESULT 11
AAW98161
ID AAW98161 standard; Protein; 140 AA.
XX AC AAW98161;
XX DT 05-JUL-1999 (first entry)
XX DE Murine Bim-L mutant T54A.
XX KW Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
KW cell cycle; mouse; cancer; autoimmune disease; therapy;
KW contraceptive; splice variant; isoform; mutant; dynein light chain;
KW cytotoxicity; agonist.
XX OS Mus musculus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX Binding-site 42..71
/note= "dynein light chain binding region, present
only in Bim-L, Bim-EL"
XX FT Misc-difference 54
/note= "Ala replaces Thr of wild-type Bim-L"
XX FT Region 94..102
/note= "BH3"
XX FT Region 123..137
/note= "hydrophobic region"
XX PN W09914321-A1.
XX PD 25-MAR-1999.
XX PF 17-SEP-1998; 98WO-AU00772.
XX PR 24-SEP-1997; 97AU-0009373.
XX PR 17-SEP-1997; 97AU-0009263.
XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX PI Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
PI Puthalakath H, Strasser A;


```

XX DR WPI; 1999-244030/20.
XX PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
XX PT treatment
XX PS Claim 25; Page -: 145pp; English.
XX CC The present sequence is a T54A mutant of the long form (L) of
XX CC murine Bim, or Bcl-2 interacting mediator of cell death (see also
XX CC AAW98155), a novel member of the Bcl-2 family. Native Bim-L is
XX CC capable of inducing cell death (apoptosis) and acts as a
XX CC 'death-ligand' for certain members of the pro-survival Bcl-2
XX CC family. Binding the dynein light chain regulates the pro-apoptotic
XX CC activity of Bim. Bim-S (see AAW98154), a splice variant which does
XX CC not bind to dynein light chain, is a much more potent killer than
XX CC either Bim-L or Bim-EL. The invention provides variants (see
XX CC AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
XX CC couple or otherwise associate with a dynein light chain owing to
XX CC amino acid addition, substitution and/or deletion within the region
XX CC which binds the dynein light chain. The variants act as agonists
XX CC of Bim function, interacting with anti-apoptotic molecules such as
XX CC Bcl-2 to prevent their functional activity, thereby promoting
XX CC apoptosis, and can be used e.g. in the treatment of cancer or to
XX CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
XX CC is expressed in germ cells, modulating its expression or activity
XX CC may be useful as a contraceptive or method of sterilization by
XX CC preventing generation of fertile sperm.
XX CC Note. The present sequence is not shown in the specification but
XX CC is derived from the murine Bim-L sequence given on page 95-96.
XX SQ Sequence 140 AA;

Query Match 54.8%; Score 584; DB 20; Length 140;
Best Local Similarity 59.5%; Pred. No. 8.1e-48;
Matches 119; Conservative 8; Mismatches 11; Indels 62; Gaps 2;

QY 1 MAKQPSDVSSCEDREGRLQQAERPPQLRPGAPTSLSQTPEQGNPEGNGEGDCSPHSP 60
Db 1 makqpsdvsscedreggqlqaerppqlrpgaptslqtpeq----- 41

QY 61 QGPLAPPASGPGPATRSPLIFMRRSLLSRSSGYFSFDTDRSPAPMSCKSTQTPSP 120
Db 42 -----drspapmscdksaqtsp 60

QY 121 COAFNHVLSAMASMQA--EPADMRPEIWTAEQLRIGDEFNAYYARRVFLNNYQAEADH 178
Db 61 cqafnhylsamasirsdqeedirpeiriaqeirrigdefnetyrrrvfandyreadh 120

QY 179 PRMVIILRLRYVRLVWRMH 198
Db 121 pqmviqlrlrfvrlvwrh 140

RESULT 12
AAW98163
ID AAW98163 standard; Protein; 140 AA.
AC AAW98163;
XX 05-JUL-1999 (first entry)
XX Murine Bim-L mutant N65S.
XX Bim-L; Bcl-2 interacting mediator of cell death; apoptosis;
XX cell cycle; mouse; cancer; autoimmune disease; therapy;
XX KW contraceptive; splice variant; isoform; mutant; dynein light chain;
XX KW cytotoxicity; agonist.
XX OS Mus musculus.
XX OS Synthetic.
XX

```

```

FH Key Location/Qualifiers
FT Binding-site 42..71
FT /note= "dynein light chain binding region, present
FT only in Bim-L, Bim-EL"
FT Misc-difference 65 /note= "Ser replaces Asn of wild-type Bim-L"
FT Region 94..102
FT /note= "BH3"
FT Region 123..137
FT /note= "hydrophobic region"
XX WO9914321-A1.
XX 25-MAR-1999.
XX 17-SEP-1998; 98WO-AU00772.
XX 24-SEP-1997; 97AU-0009373.
XX 17-SEP-1997; 97AU-0009263.
XX (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX Adams J, Cory S, Huang DCS, O'Connor L, O'Reilly L;
XX Puthalakath H, Strasser A;
XX WPI; 1999-244030/20.
XX New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
XX treatment
XX Claim 25; Page -: 145pp; English.
XX The present sequence is a N65S mutant of the long form (L) of
XX murine Bim, or Bcl-2 interacting mediator of cell death (see also
XX AAW98155), a novel member of the Bcl-2 family. Native Bim-L is
XX capable of inducing cell death (apoptosis) and acts as a
XX 'death-ligand' for certain members of the pro-survival Bcl-2
XX family. Binding the dynein light chain regulates the pro-apoptotic
XX activity of Bim. Bim-S (see AAW98154), a splice variant which does
XX not bind to dynein light chain, is a much more potent killer than
XX either Bim-L or Bim-EL. The invention provides variants (see
XX AAW98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
XX couple or otherwise associate with a dynein light chain owing to
XX amino acid addition, substitution and/or deletion within the region
XX which binds the dynein light chain. The variants act as agonists
XX of Bim function, interacting with anti-apoptotic molecules such as
XX Bcl-2 to prevent their functional activity, thereby promoting
XX apoptosis, and can be used e.g. in the treatment of cancer or to
XX delete autoreactive lymphocytes in autoimmune diseases. Since Bim
XX is expressed in germ cells, modulating its expression or activity
XX may be useful as a contraceptive or method of sterilization by
XX preventing generation of fertile sperm.
XX Note. The present sequence is not shown in the specification but
XX is derived from the murine Bim-L sequence given on page 95-96.
XX SQ Sequence 140 AA;

Query Match 54.8%; Score 584; DB 20; Length 140;
Best Local Similarity 59.5%; Pred. No. 8.1e-48;
Matches 119; Conservative 9; Mismatches 10; Indels 62; Gaps 2;

QY 1 MAKQPSDVSSCEDREGRLQQAERPPQLRPGAPTSLSQTPEQGNPEGNGEGDCSPHSP 60
Db 1 makqpsdvsscedreggqlqaerppqlrpgaptslqtpeq----- 41

QY 61 QGPLAPPASGPGPATRSPLIFMRRSLLSRSSGYFSFDTDRSPAPMSCKSTQTPSP 120
Db 42 -----drspapmscdksaqtsp 60

QY 121 COAFNHVLSAMASMQA--EPADMRPEIWTAEQLRIGDEFNAYYARRVFLNNYQAEADH 178
Db 61 cqafnhylsamasirsdqeedirpeiriaqeirrigdefnetyrrrvfandyreadh 120

```


DR WPI: 1999-244030/20.
XX
PT New isolated member of the Bcl-2 family, Bim used in, e.g. cancer
PT treatment
XX
XX
PS Claim 25; Page -: 145pp; English.
XX
CC The present sequence is a D51G mutant of the long form (L) of
CC murine Bim, or Bcl-2 interacting mediator of cell death (see also
CC AA98155), a novel member of the Bcl-2 family. Native Bim-L is
CC capable of inducing cell death (apoptosis) and acts as a
CC 'death-ligand' for certain members of the pro-survival Bcl-2
CC family. Binding the dynein light chain regulates the pro-apoptotic
CC activity of Bim. Bim-S (see AA98154), a splice variant which does
CC not bind to dynein light chain, is a much more potent killer than
CC either Bim-L or Bim-EL. The invention provides variants (see
CC AA98159-68) of murine and human Bim-L and Bim-EL that cannot bind,
CC couple or otherwise associate with a dynein light chain owing to
CC amino acid addition, substitution and/or deletion within the region
CC which binds the dynein light chain. The variants act as agonists
CC of Bim function, interacting with anti-apoptotic molecules such as
CC Bcl-2 to prevent their functional activity, thereby promoting
CC apoptosis, and can be used e.g. in the treatment of cancer or to
CC delete autoreactive lymphocytes in autoimmune diseases. Since Bim
CC is expressed in germ cells, modulating its expression or activity
CC may be useful as a contraceptive or method of sterilization by
CC preventing generation of fertile sperm.
CC Note. The present sequence is not shown in the specification but
CC is derived from the murine Bim-L sequence given on page 95-96.
XX
SQ Sequence 140 AA;

Query Match 54.6%; Score 582; DB 20; Length 140;
Best Local Similarity 59.5%; Pred. No. 1.3e-47;
Matches 119; Conservative 8; Mismatches 11; Indels 62; Gaps 2;

QY 1 MAKQPSDVSSCDREGRLQPAERPPQLRPGAPTSLSQTEPQGNHGGEGDSCPHGSP 60
Db 1 makqpsdvssecdreggqlpaerppqlrpgaptslqtepq----- 41

QY 61 QGPLAPPASGPGPATRSPLFIFMRRSLLSRSSSGYFSDTRSPAPMSCDKSTQTPSP 120
Db 42 -----drspapmscgkstqtpsp 60

QY 121 COAFNHYLSAMASMRQA--EPADMRPEIWTQELRIGDEFNAYYARRVFLNNYQAEDEH 178
Db 61 cqaafnhyisamasirsdqeedlrpeirilaagelrrigdefnetyrrrvfrandyreaedh 120

QY 179 PRMVILRLRYIVRLVWRMH 198
Db 121 pqmviqlrlfrflwrwh 140

RESULT 15
ABG01163
ID ABG01163 standard; Protein; 92 AA.
XX
AC ABG01163;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #1154.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX

PF 30-MAR-2001; 2001WO-US08631.
XX
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang VT;
PI
XX
XX WPI: 2001-639362/73.
DR N-PSDB; AAS65350.
DR
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID No 31522; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 92 AA;

Query Match 46.7%; Score 497; DB 22; Length 92;
Best Local Similarity 100.0%; Pred. No. 9e-40;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQPSDVSSCDREGRLQPAERPPQLRPGAPTSLSQTEPQGNHGGEGDSCPHGSP 60
Db 1 makqpsdvssecdreggqlpaerppqlrpgaptslqtepqgnhggcdscphgsp 60

QY 61 QGPLAPPASGPGPATRSPLFIFMRRSLLSR 92
Db 61 qgplappasgpgpatrsplfifmrrsllsrs 92

Search completed: August 16, 2002, 10:00:59
Job time: 262 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 09:57:22 ; Search time 13.18 Seconds
(without alignments)
366.940 Million cell updates/sec

Title: US-09-508-832-10
Perfect score: 1065
Sequence: 1 MAKQPSDVSSECDREGRLQ.....PRMVLRLRYIVRLVRMH 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111.5	10.5	886	US-08-474-379C-65	Sequence 65, Appl
2	111.5	10.5	886	US-09-146-249A-65	Sequence 65, Appl
3	111.5	10.5	886	US-08-206-188B-65	Sequence 65, Appl
4	108.5	10.2	885	US-08-577-492-33	Sequence 33, Appl
5	108.5	10.2	885	US-09-079-630-33	Sequence 33, Appl
6	99	9.3	228	US-09-219-849-38	Sequence 38, Appl
7	99	9.3	1057	US-08-931-820-4	Sequence 4, Appl
8	99	9.3	1078	US-08-963-825-21	Sequence 21, Appl
9	99	9.3	1078	US-09-570-573-21	Sequence 21, Appl
10	99	9.3	1078	US-09-548-608-21	Sequence 21, Appl
11	94.5	8.9	280	US-09-247-155-178	Sequence 178, App
12	94.5	8.9	1185	US-09-041-886-23	Sequence 23, Appl
13	94	8.8	580	US-08-906-865-1	Sequence 1, Appl
14	94	8.8	627	US-08-466-589-6	Sequence 6, Appl
15	94	8.8	627	US-08-700-636-6	Sequence 6, Appl
16	94	8.8	627	US-08-467-574-6	Sequence 6, Appl
17	94	8.8	627	US-09-217-345-6	Sequence 6, Appl
18	93.5	8.8	434	US-08-710-249-4	Sequence 4, Appl
19	93.5	8.8	434	US-09-220-157A-4	Sequence 4, Appl
20	93.5	8.8	902	US-08-396-479B-6	Sequence 6, Appl
21	93.5	8.8	902	US-08-818-823-6	Sequence 6, Appl
22	92.5	8.7	245	US-09-140-804-4	Sequence 4, Appl
23	92.5	8.7	829	US-08-642-255-132	Sequence 132, App
24	92.5	8.7	829	US-08-397-633A-53	Sequence 53, Appl
25	92	8.6	761	US-08-707-237A-84	Sequence 84, Appl
26	92	8.6	762	US-08-642-255-114	Sequence 114, App
27	92	8.6	762	US-08-642-255-120	Sequence 120, App

28	92	8.6	762	1	US-08-397-633A-26	Sequence 26, Appl
29	92	8.6	762	1	US-08-397-633A-31	Sequence 31, Appl
30	91.5	8.6	1319	2	US-08-290-731C-2	Sequence 2, Appl
31	91.5	8.6	1336	2	US-08-290-731C-6	Sequence 6, Appl
32	90	8.5	1323	1	US-08-026-138E-4	Sequence 4, Appl
33	90	8.5	1706	2	US-08-459-568-2	Sequence 2, Appl
34	90	8.5	1706	2	US-08-399-411-2	Sequence 2, Appl
35	90	8.5	1706	3	US-08-516-859A-2	Sequence 2, Appl
36	89.5	8.4	123	4	US-09-247-155-91	Sequence 91, Appl
37	89	8.4	1719	2	US-08-459-568-4	Sequence 4, Appl
38	89	8.4	1719	2	US-08-399-411-4	Sequence 4, Appl
39	89	8.4	1719	3	US-08-516-859A-4	Sequence 4, Appl
40	88.5	8.3	1333	3	US-09-356-952-2	Sequence 2, Appl
41	88	8.3	357	1	US-07-609-716-66	Sequence 66, Appl
42	88	8.3	357	1	US-08-642-255-33	Sequence 33, Appl
43	88	8.3	357	4	US-08-475-411A-66	Sequence 66, Appl
44	88	8.3	357	4	US-08-478-029A-66	Sequence 66, Appl
45	87.5	8.2	2337	3	US-08-713-118-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-474-379C-65
; Sequence 65, Application US/08474379C
; Patent No. 5977305
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: CLONING BY COMPLEMENTATION AND RELATED
; TITLE OF INVENTION: PROCESSES
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,379C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206,188
; FILING DATE: 01-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/688,352
; FILING DATE: 19-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 27866/32771
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 886 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-474-379C-65

Query Match 10.5%; Score 111.5; DB 2; Length 886;
Best Local Similarity 27.5%; Pred. No. 0.0046;
Matches 47; Conservative 18; Mismatches 67; Indels 39; Gaps 7;
QY 6 SDVSECDREGROLQPAERPQL----RPGAPTSLOTPEQ-----GNPEGNHGG----- 50
Db 49 SDSAERAERERQPHRIERADAMTSDRGLTRTMSWPSFHGTGTGSGGAGGSSRRF 108
QY 51 EGDSCPHGSP-QGPLAPPSPGPATRSPLFIFMRRSSLLSRSSSGYFSDTDRSPAPMS 109
Db 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162
QY 110 CDKSTQTPSPCCAFNHYLSAMASMRQAEPAADMRPEIWIQAELRRIGDEFN 160
Db 163 RN-----SSVTSEAHAEIDLIVTPFAQVLASLRSVRNSFNFS 196
RESULT 2
US-09-146-249A-65
; Sequence 65, Application US/09146249A
; Patent No. 6069240
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,249A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 886 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-146-249A-65
Query Match 10.5%; Score 111.5; DB 3; Length 886;
Best Local Similarity 27.5%; Pred. No. 0.0046;
Matches 47; Conservative 18; Mismatches 67; Indels 39; Gaps 7;
QY 6 SDVSECDREGROLQPAERPQL----RPGAPTSLOTPEQ-----GNPEGNHGG----- 50
Db 49 SDSAERAERERQPHRIERADAMTSDRGLTRTMSWPSFHGTGTGSGGAGGSSRRF 108
QY 51 EGDSCPHGSP-QGPLAPPSPGPATRSPLFIFMRRSSLLSRSSSGYFSDTDRSPAPMS 109
Db 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162
QY 110 CDKSTQTPSPCCAFNHYLSAMASMRQAEPAADMRPEIWIQAELRRIGDEFN 160
Db 163 RN-----SSVTSEAHAEIDLIVTPFAQVLASLRSVRNSFNFS 196

Db 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162
QY 110 CDKSTQTPSPCCAFNHYLSAMASMRQAEPAADMRPEIWIQAELRRIGDEFN 160
Db 163 RN-----SSVTSEAHAEIDLIVTPFAQVLASLRSVRNSFNFS 196
RESULT 3
US-08-206-188B-65
; Sequence 65, Application US/08206188B
; Patent No. 6100025
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 84
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,188B
; FILING DATE: 01-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/511,715
; FILING DATE: 20-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 886 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-206-188B-65
Query Match 10.5%; Score 111.5; DB 3; Length 886;
Best Local Similarity 27.5%; Pred. No. 0.0046;
Matches 47; Conservative 18; Mismatches 67; Indels 39; Gaps 7;
QY 6 SDVSECDREGROLQPAERPQL----RPGAPTSLOTPEQ-----GNPEGNHGG----- 50
Db 49 SDSAERAERERQPHRIERADAMTSDRGLTRTMSWPSFHGTGTGSGGAGGSSRRF 108
QY 51 EGDSCPHGSP-QGPLAPPSPGPATRSPLFIFMRRSSLLSRSSSGYFSDTDRSPAPMS 109
Db 109 EAENGTPSPGRPLDSQASPG-LVLHAGAATSQRRESFLYRSDSY-----DMSPKTMS 162
QY 110 CDKSTQTPSPCCAFNHYLSAMASMRQAEPAADMRPEIWIQAELRRIGDEFN 160
Db 163 RN-----SSVTSEAHAEIDLIVTPFAQVLASLRSVRNSFNFS 196
RESULT 4
US-08-577-492-33
; Sequence 33, Application US/08577492
; Patent No. 5851784

```

RESULT      5
US-09-079-630-33
; Sequence 33, Application US/09079630
; Patent No. 6291199
;
; GENERAL INFORMATION:
; APPLICANT: Owens, Raymond John
; APPLICANT: Perry, Martin John
; APPLICANT: Lumb, Simon Mark
;
; TITLE OF INVENTION: HUMAN PHOSPHODIESTERASE TYPE IVC, AND
;
; TITLE OF INVENTION: ITS PRODUCTION AND USE
;
; NUMBER OF SEQUENCES: 40
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock,
; Kurtz, Mackiewicz & No. 6291199rlis

```

```

RESULT      6
US-09-219-849-38
US-09-219-849-38, Application US/09219849
; Patent No. 6150081
; GENERAL INFORMATION:
; APPLICANT: VAN HEERDE, GEORGE V.
; APPLICANT: VAN RIJN, ALEXIS C.
; APPLICANT: BOWSTRAN, JAN B.
; APPLICANT: DE WOLF, FREDRIK A.
; APPLICANT: MOOBROEK, ANDREAS
; APPLICANT: WERTEN, MARC W.T.
; APPLICANT: WIND, RICHEL D.
; APPLICANT: VAN DEN BOSCH, TANJA J.
; TITLE OF INVENTION: SILVER HALIDE
; TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHY
; TITLE OF INVENTION: PREPARATION THIN FILMS
; FILE REFERENCE: 2728-2
; CURRENT APPLICATION NUMBER: US/09/219-849-38
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 50

```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid sequence
US-09-219-849-38

Query Match          9.3%; Score 99; DB 4; Length 228;
Best Local Similarity 35.4%; Pred. No. 0.015;
Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;

QY 5 PSDVSSCDREGROLPAPER---PQLRPGAP-TSLQTEPGNP-----EGNHGEG 52
Db 66 PQGKSGSGKPGASHNGERPPGQGLPQGTAGPGRDGNPGSDGQPGRDGSGGKG 125
QY 53 DSCPHGSPQGPLAP--PASPGP 72
Db 126 DRGENSGPAGPAGHPGPGPV 147

RESULT 7
US-08-931-820-4
; Sequence 4, Application US/08931820
; Patent No. 6010863
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Assay for collagen degradation
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,820
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 96202596.1
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1057 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Collagen type III
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1055
; OTHER INFORMATION: /label= Modified
; OTHER INFORMATION: /note= "Ala may be Pro"
US-08-931-820-4

Query Match          9.3%; Score 99; DB 3; Length 1057;
Best Local Similarity 27.8%; Pred. No. 0.11;
Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;

QY 5 PSDVSSCDREGROLPAPERPPQLRPGAPTSLQTEPGNP-----EGNHGEG 52
Db 825 PQGKSGSGKPGANGLSGERPPGQGLPGLAGTAGPGRDGNPGSDGLPGRDGSPPGKG 884
QY 53 DSCPHGSPQGPLAP--PASPGPFPATRSPLFTFMRRSSLSRSSSGYFSDTDRSPAPMSC 110
Db 126 DRGENSGPAGPAGHPGPGPV 147

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 111
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid sequence
US-09-219-849-38

Query Match          9.3%; Score 99; DB 3; Length 1078;
Best Local Similarity 27.8%; Pred. No. 0.11;
Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;

QY 5 PSDVSSCDREGROLPAPERPPQLRPGAPTSLQTEPGNP-----EGNHGEG 52
Db 826 PQGKSGSGKPGANGLSGERPPGQGLPGLAGTAGPGRDGNPGSDGLPGRDGSPPGKG 885
QY 53 DSCPHGSPQGPLAP--PASPGPFPATRSPLFTFMRRSSLSRSSSGYFSDTDRSPAPMSC 110
Db 886 DRGENSGPAGPAGHPGPGPV 147
QY 111 DKSTQTPSPCQA 123

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 123
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid sequence
US-09-219-849-38

Query Match          9.3%; Score 99; DB 3; Length 1078;
Best Local Similarity 27.8%; Pred. No. 0.11;
Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;

QY 5 PSDVSSCDREGROLPAPERPPQLRPGAPTSLQTEPGNP-----EGNHGEG 52
Db 826 PQGKSGSGKPGANGLSGERPPGQGLPGLAGTAGPGRDGNPGSDGLPGRDGSPPGKG 885
QY 53 DSCPHGSPQGPLAP--PASPGPFPATRSPLFTFMRRSSLSRSSSGYFSDTDRSPAPMSC 110
Db 886 DRGENSGPAGPAGHPGPGPV 147
QY 111 DKSTQTPSPCQA 123
```


[illegible]

1 APPLICANT: BALPOLLO, MICHAEL M.
 2
 3 TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
 4
 5 TITLE OF INVENTION: RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
 6
 7 NUMBER OF SEQUENCES: 12
 8
 9 CORRESPONDENCE ADDRESS:
 10 ADDRESS: Pretty, Schroeder, Brueggemann & Clark
 11 STREET: 444 South Flower Street, Suite 2000
 12 CITY: Los Angeles
 13 STATE: CA
 14 COUNTRY: USA
 15
 16 ZIP: 90071
 17
 18 COMPUTER READABLE FORM:

```

Query Match      8.8%; Score 94; DB 2; Length 627;
Best Local Similarity 25.6%; Pred. No. 0.18;
Matches 32; Conservative 14; Mismatches 45; Indels 34; Gaps

Qy      1 MAKOPDVSSECOREGQLOPAERPOLRPGATPSTLQTEPQGNPEGNHGEGDSCPHGSP 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      357 LMKRPSVVKNCRRLLIESMHKMASAPRFW-----EPEGPATSGTQSLHPSPPF 408

Qy      61 QGPLAPASGPGPATRSPGLIFMRSSLRSRSGVFSFDTDRSP--APMSCDKSTQTPS 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      409 CVPLLDVPAEPPG-SCKSP-----SDLPQPKPLAEAKSDSPHS 445

```

Qy 119 P-PCQ 122
 | 11:
Db 446 PGPCR 450

Search completed: August 16, 2002, 10:01:20
Job time: 238 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Query Match 9.7%; Score 103.5; DB 2; Length 556;
Best Local Similarity 27.8%; Pred. No. 0.71;
Matches 42; Conservative 17; Mismatches 51; Indels 41; Gaps 8;

QY 1 MAKQPSD-----VSECDREGROLOPAERPPQLRPGAPTSLOTPEQGNPEGN-----HG 49
DB LKKNPNERFPSABAMRDECLRVAASFQAA--PPSIVPGAQTS-----SGAGVGSVAFPPV 320

QY 50 GEGDSCPHGSPGQPLAPPASPGPF-----ATRSPLFTFMRRSSLLSRSSSGYFSFDTD 102
DB 321 GQGTAPAGVQPTQPTLSPGPNPYGTAPAAHSPAYGYPOQ-----AGY-----QT 368

QY 103 RSPAPMSCDKSTQTPSPPCQAFNHYLSAMAS 133
DB 369 PAPAPYAQQQAATPPP-----YNLTPSAOGS 395

RESULT 6
T36502
serine/threonine protein kinase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T36502
R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21608
A:Accession: T36502
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-536 <SAU>
A:Cross-references: EMBL:AL096822; PIDN:CAB46944.1; GSPDB:GN00070; SCOEDB:SCGD3.22
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: pksC; SCOEDB:SCGD3.22
C:Superfamily: Synechocystis sp. protein kinase, 58K; protein kinase homology

Query Match 9.7%; Score 103.5; DB 2; Length 556;
Best Local Similarity 27.8%; Pred. No. 0.71;
Matches 42; Conservative 17; Mismatches 51; Indels 41; Gaps 8;

QY 1 MAKQPSD-----VSECDREGROLOPAERPPQLRPGAPTSLOTPEQGNPEGN-----HG 49
DB LKKNPNERFPSABAMRDECLRVAASFQAA--PPSIVPGAQTS-----SGAGVGSVAFPPV 320

QY 50 GEGDSCPHGSPGQPLAPPASPGPF-----ATRSPLFTFMRRSSLLSRSSSGYFSFDTD 102
DB 321 GQGTAPAGVQPTQPTLSPGPNPYGTAPAAHSPAYGYPOQ-----AGY-----QT 368

QY 103 RSPAPMSCDKSTQTPSPPCQAFNHYLSAMAS 133
DB 369 PAPAPYAQQQAATPPP-----YNLTPSAOGS 395

RESULT 7
I48314
homeotic protein CDP - mouse
N:Alternate names: CCAAT displacement protein; homeotic protein Cux
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: I48314
R:Valarche, I.; Tissier-Seta, J.P.; Hirsch, M.R.; Martinez, S.; Goridis, C.; Brunet, J.E.
Development 119, 881-896, 1993
A:Title: The mouse homeodomain protein phox2 regulates Ncam promoter activity in concert
with the mouse homeodomain protein Phox2b
A:Reference number: I48314; MUID:194244481
A:Accession: I48314
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1332 <RES>
A:Cross-references: EMBL:X75013; NID:g402589; PIDN:CAA52922.1; PID:g402590
C:Genetics:
A:Gene: Cux
C:Superfamily: homeotic protein CDP; cut repeat homology; homeobox homology

C:Keywords: DNA binding; duplication; homeobox; nucleus; transcription regulation
F:366-438/Domain: cut repeat homology <CU1>
F:755-827/Domain: cut repeat homology <CU2>
F:938-1010/Domain: cut repeat homology <CU3>
F:1057-1113/Domain: homeobox homology <HOX>

Query Match 9.6%; Score 102.5; DB 1; Length 1332;
Best Local Similarity 23.7%; Pred. No. 2.2;
Matches 54; Conservative 25; Mismatches 76; Indels 73; Gaps 9;

QY 6 SDVSSCEDREGROLOPAERPPQLRPGAPTSLOT---EPQGNPEGNHGEGDSCPHGSPQG 62
DB 219 SDLSGSAARRKGRQDPESRRPCPLPASPPQLPRNTGQVNTNGTH-----HFSPAG 270

QY 63 -----PLAPPASPGPFATRSPLFTFMRRS-----SLLSRSSSGYFSF 100
DB 271 LSQDFTSSNLASPLASTGKFAFALNSLLQRLMQSFYSKAMQEAAGSTSTFTGTPYSTN 330

QY 101 TDRSPAPM--SCDKSTQTPSP-----COAFN 125
DB 331 SISPSPLOQSPDVNGMAPSPQSSESAGSISEGEIDTAIAQVQKQKQLIKHNIGQRIFG 390

QY 126 HYLSAMASMRQAPADMRPPIAIAQELRLRGDEFNAYARRVFLNNYQ 173
DB 391 HYVLGLSQSGVSEIL-ARPKPWNKLTVR--GKE--PFHKMKQFLSDRQ 433

RESULT 8
CGBEHS
collagen-like protein - salmairine herpesvirus 1 (strain 484-77)
C:Species: salmairine herpesvirus 1
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Feb-1995
C:Accession: A36770
R:Geck, P.; Whitaker, S.A.; Medveczky, M.M.; Medveczky, P.G.
J. Virol. 64, 3509-3515, 1990
A:Title: Expression of collagen-like sequences by a tumor virus, Herpesvirus salmairi.
A:Reference number: A36770; MUID:90279084
A:Accession: A36770
A:Molecule type: DNA
A:Residues: 1-99 <GEC>
A:Cross-references: EMBL:M31965
C:Superfamily: squirrel monkey herpesvirus collagen-like protein
F:15-70/Region: collagen-like

Query Match 9.4%; Score 100; DB 1; Length 99;
Best Local Similarity 36.6%; Pred. No. 0.21;
Matches 30; Conservative 6; Mismatches 32; Indels 14; Gaps 4;

QY 1 MAKQPS--DVSSCEDREGROLOPAERPPQLRPGAPTSLOTPEQG--NPEGNHGEGDSCP 56
DB 1 MASEPNLRYPTETGDRGPGPGPPGPPG-----PQGPQPGPPGPPGPPG 52

QY 57 HG--SPQGPLAPPASPGPFATR 76
DB 53 QGPPGPGPPGPPGPPGPPSDSR 74

RESULT 9
JC4021
nicotinic acetylcholine receptor alpha-4 chain, neuronal - human
C:Species: Homo sapiens (man)
C>Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 20-Aug-1999
C:Accession: JC4021; S55471
R:Monteggia, L.M.; Gopalakrishnan, M.; Touma, E.; Idler, K.B.; Nash, N.; Arneric, S.P.
Gene 155, 189-193, 1995
A:Title: Cloning and transient expression of genes encoding the human alpha 4 and bet
A:Reference number: JC4021; MUID:95237608
A:Accession: JC4021
A:Molecule type: mRNA
A:Residues: 1-627 <MON>
A:Cross-references: GB:IJ35901; NID:g755647; PIDN:AAA64743.1; PID:g755648

R;Mamalaki, A.; Remoundos, M.; Tzartos, S.
submitted to the EMBL Data Library, May 1995
A;Description: Molecular cloning of human neuronal nicotinic acetylcholine receptor 4-13

A;Reference number: S55471
A;Accession: S55471
A;Status: preliminary
A:Molecule type: mRNA
A;Residues: 26-627 <NAM>
A;Cross-references: EMBL:X87629; NID:g854158; PIDN:CAA60959.1; PID:g854159
C;Genetics:
A;Gene: GDB:CHRNA4
A;Cross-references: GDB:I28169; OMIM:118504
A;Map position: 20q13.2-20q13.3
C;Superfamily: acetylcholine receptor
C;Keywords: ion channel; neurotransmitter receptor; postsynaptic membrane; transmembrane

Query Match 9.4%; Score 100; DB 2; Length 627;
Best Local Similarity 26.4%; Pred. No. 1.5;
Matches 33; Conservative 15; Mismatches 43; Indels 34; Gaps 5;

QY 1 MAKOPDSVSSECDREGRLQLOPAERPPLRPGAPTSLQTEPOGNGEGHGSDSCPHGSP 60
: l : l : l : l : l : l : l : l : l : l : l : l : l : l : l : l : l : l :
Db 357 LMKRPVVKDKCRLLIESMHKMASAPFWP-----EPEGEPPATSGTSLHPPPSF 408

QY 61 QGPLAPPASPFGPATFRPLFLIFMRSSLLSRSSGYFSFDTRSP--APMSCDKSTOTPS 118
: l : l : l : l : l : l : l : l : l : l : l : l : l : l : l : l : l : l :
Db 409 CVPLDVPAPGP-SCKSP-----SDQLPPQQPLEAEKASHPHS 445

QY 119 P-PQG 122
: l : l :
Db 446 PGPCR 450

RESULT 10
S41067
collagen alpha 1(III) chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Aug-1999
C;Accession: S41067; A29905; S31924
R;Glumoff, V.; Maekelae, J.K.; Vuorio, E.
Biochim. Biophys. Acta 1217, 41-48, 1994
A;Title: Cloning of cDNA for rat pro alpha-1(III) collagen mRNA. Different expression pattern
A;Reference number: S41067; MUID:94114571
A;Accession: S41067
A;Status: preliminary
A:Molecule type: mRNA
A;Residues: 1-636 <GLU>
A;Cross-references: EMBL:X70369; NID:g57915; PIDN:CAA49832.1; PID:g57916
R;Frankel, F.R.; Hsu, C.Y.J.; Myers, J.C.; Lin, E.; Lyttle, C.R.; Komm, B.; Mohr, K.
DNA 7, 347-354, 1988
A;Title: Regulation of alpha-2 (I), alpha-1 (III), and alpha-2(V) collagen mRNAs by estradiol
A;Reference number: A29905; MUID:88296083
A;Accession: A29905
A:Molecule type: mRNA
A;Residues: 308-482 <FRA>
A;Cross-references: GB:M21354; NID:g203500; PIDN:AAA40942.1; PID:g203501
R;Glumoff, V.; Maekelae, J.K.; Vuorio, E.
submitted to the EMBL Data Library, February 1993
A;Reference number: S31924
A;Accession: S31924
A;Status: preliminary
A:Molecule type: mRNA
A;Residues: 2-636 <GL2>
A;Cross-references: EMBL:X70369
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F;408-636/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 9.3%; Score 99; DB 2; Length 636;
Best Local Similarity 35.4%; Pred. No. 1.9;
Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;

submitted to the EMBL Data Library, February 1989
 A:Reference number: S05272
 A:Accession: S05272

A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-1240, 'V', 1242-1466 <RNC>
 A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1.; PID:g30058

R:Ala-Kokko, L.; Kontusaari, S.; Baldwin, C.T.; Kulvaneniemi, H.; Prockop, D.J.
 Biochem. J. 260, 509-516, 1989

A:Title: Structure of cDNA clones coding for the entire prepro-alpha(III) chain of human procollagen.

A:Reference number: S04642; MUID:89350838
 A:Accession: S04642

A:Molecule type: mRNA
 A:Residues: 1-1196 <ALA>

A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1.; PID:g30058
 A:Note: the complete sequence is not shown

R:Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.

Gene 78, 255-265, 1989

A:Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene (COL3A1).

A:Reference number: P80011; MUID:89378752
 A:Accession: P80011

A:Molecule type: DNA
 A:Residues: 1-176 <BEN>

A:Cross-references: GB:M26939; NID:g180813; PIDN:AAA52040.1.; PID:g180814
 R:Toman, P.D.; Ricca, G.A.; de Crombrughe, B.

Nucleic Acids Res. 16, 7201, 1988

A:Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human procollagen type I.

A:Reference number: S01726; MUID:89303360
 A:Accession: S01726

A:Molecule type: mRNA
 A:Residues: 1-170 <TON>

A:Cross-references: EMBL:X07240; NID:g30060; PIDN:CAA30229.1.; PID:g30061
 A:Note: the authors translated the codon CAG for residue 154 as His

R:Janeczko, R.A.; Ramirez, F.

Nucleic Acids Res. 17, 6742, 1989

A:Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen.
 A:Reference number: S04887; MUID:89386015
 A:Accession: S04887

A:Molecule type: mRNA
 A:Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-634,

A:Cross-references: EMBL:X15332; NID:g29545; PIDN:CAA33387.1.; PID:g9930045
 A:Note: the authors' translation of residues 905-932 is inconsistent with the nucleotide

R:Seyer, J.M.; Kang, A.H.

Biochemistry 16, 1158-1164, 1977

A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptide (alpha-1(I) procollagen).

A:Reference number: A90399; MUID:77134724
 A:Accession: A90399

A:Molecule type: protein
 A:Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEY1>

A:Experimental source: liver

A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose

R:Seyer, J.M.

submitted to the Atlas, December 1977

A:Reference number: A94562
 A:Accession: A94562

A:Molecule type: protein
 A:Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>

A:Experimental source: liver

A:Note: author submitted corrections to A90399

R:Milwicz, D.M.; Wit, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.

Am. J. Hum. Genet. 53, 62-70, 1993

A:Title: Parental somatic and germ-line mosaicism for a multigene deletion with unusual frequency.

A:Reference number: I51868; MUID:93304430
 A:Accession: I51868

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 186-194 <MIL>

A:Cross-references: GB:S62925; NID:g386425; PIDN:AAD13937.1.; PID:g4261637
 R:Chiodo, A.A.; Sillence, D.O.; Cole, W.G.; Bateman, J.F.

Biochem. J. 311, 939-943, 1995

A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3A1 gene.

A:Reference number: S59511; MUID:96067614
 A:Accession: S59511

A:Molecule type: mRNA
 A:Residues: 302-423 <CHI>

A:Cross-references: GB:S79877; NID:g1195576; PIDN:AAB35615.1.; PID:g1195577
 R:Seyer, J.M.; Kang, A.H.

Biochemistry 17, 3404-3411, 1978

A:Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBR

A:Reference number: A90414; MUID:79000343
 A:Accession: A90414

A:Molecule type: protein
 A:Residues: 399-675, 'N', 677-727 <SEY3>

A:Experimental source: liver

R:Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.

J. Biol. Chem. 266, 5256-5259, 1991

A:Title: G to T transition at position +5 of a splice donor site causes skipping of exon 17 in the COL3A1 gene.

A:Reference number: I55349; MUID:91161621
 A:Accession: I55349

A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 537-605 <LEE>

A:Cross-references: GB:M59312; NID:g180815; PIDN:AAA52041.1.; PID:g180816
 R:Seyer, J.M.; Mainardi, C.; Kang, A.H.

Biochemistry 19, 1583-1589, 1980

A:Title: Covalent structure of collagen: amino acid sequence of alpha 1 (III)-CB5 from the COL3A1 gene.

A:Reference number: A90438; MUID:80198282
 A:Accession: A90438

A:Molecule type: protein
 A:Residues: 728-895, 'A', 897-964 <SEY4>

A:Experimental source: liver
 R:Chiodo, W.G.; Chiodo, A.A.; Lamande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; C

J. Biol. Chem. 265, 17070-17077, 1990

A:Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping in the COL3A1 gene.

A:Reference number: A38303; MUID:91009133
 A:Accession: A38303

A:Molecule type: mRNA
 A:Residues: 861-1015 <COL>

A:Cross-references: GB:J05617; GB:M55603; GB:M59227; NID:g180878; PIDN:AAB59383.1.; PI
 A:Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos

R:ManKoo, B.S.; Dalgleish, R.

Nucleic Acids Res. 16, 2337, 1988

A:Title: Human pro alpha1(I) collagen: cDNA sequence for the 3' end.

A:Reference number: S02119; MUID:88189827
 A:Accession: S02119

A:Status: translation not shown

A:Molecule type: mRNA
 A:Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>

A:Cross-references: EMBL:X06700; NID:g30053; PIDN:CAA29886.1.; PID:g30054
 R:Seyer, J.M.; Kang, A.H.

Biochemistry 20, 2621-2627, 1981

A:Title: Covalent structure of collagen: amino acid sequence of alpha 1 (III)-CB9 from the COL3A1 gene.

A:Reference number: A90446; MUID:81508139
 A:Accession: A90446

A:Molecule type: protein
 A:Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 11

A:Experimental source: liver
 R:Loidl, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.;
 Nucleic Acids Res. 12, 9383-9394, 1984

A:Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollagen.

A:Reference number: A93551; MUID:85087944
 A:Accession: A93551

A:Molecule type: mRNA
 A:Residues: 1065-1155, 'P', 1157-1466 <LOI>

A:Cross-references: EMBL:X01655; EMBL:X01742; NID:g29584; PIDN:CAA25821.1
 R:Wiskuln, M.; Dalgleish, R.; Kluge-Beckerman, B.; Rennard, S.I.; Tolstoshev, P.; Br

Biochemistry 25, 1408-1413, 1986

A:Title: Human type III collagen gene expression is coordinately modulated with the type III procollagen gene.

A:Reference number: I52393; MUID:86187804
 A:Accession: I52393

A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1161-1200 <WIS>

A:Cross-references: GB:M13146; NID:g180415; PIDN:AAA52003.1.; PID:g180416

R;Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.M.; Myers, J.C.
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985
A;Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm
C;Reference number: I59025; MUID:85216505
A;Accession: I79359
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1165-1196 <EMA>
A;Cross-references: GB:M1134; NID:g180417; PIDN:AAA52004.1; PID:g180418
R;Chu, M.L.; Weil, D.; de Wet, W.; Bernard, M.; Sippola, M.; Ramirez, F.
J. Biol. Chem. 260, 4357-4363, 1985
A;Title: Isolation of cDNA and genomic clones encoding human pro-alpha(III) collagen. F
A;Reference number: A92516; MUID:85157600
A;Accession: A92516
A;Molecule type: DNA
A;Residues: 1176-1240, 'V' 1242-1356, 'P' 1358-1466 <CHU>
A;Cross-references: GB:M10615; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB
A;Experimental source: liver
A;Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given
ation
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit
3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently
C;Genetics:
A;Gene: GDB:COL3A1
A;Cross-references: GDB:I118729; OMIM:120180
A;Map position: 2q31-2q31
A;Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1275/3; 1418/3
A;Note: the list of introns is incomplete; defects in this gene can result in Ehlers-Dan
C;Complex: type III collagen is a homotrimer of monomers initially linked by disulfide
er of their length, is formed with desmosine cross-links made from lysine and allylsine
C;Function:
A;Description: structural component of extracellular fibrous polymer that maintains inte
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C;Keywords: coiled coil; Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hyd
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-153/Domain: amino-terminal propeptide #status predicted <PRO>
F:31-91/Domain: von Willebrand factor type C repeat homology <VWC>
F:154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>
F:154-167/Region: amino-terminal nonhelical telopeptide
F:168-1196/Region: helical
F:1091-1093/Region: cell attachment (R-G-D) motif
F:1197-1221/Region: carboxyl-terminal nonhelical telopeptide
F:1222-1466/Domain: carboxyl-terminal propeptide #status predicted <CPR>
F:1238-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted
F:154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:161,1212/Modified site: allylsine (Lys) #status predicted
F:263,284,860,977,1106/Modified site: 5-hydroxylysine (Lys) #status experimental
F:263/Binding site: carbohydurate (Lys) (covalent) #status experimental
F:584,1094/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental
F:948-949/Cleavage site: Gly-Ile (collagenase) #status experimental
F:1106/Binding site: carbohydurate (Lys) (covalent) #status predicted

Query Match 9.3%; Score 99; DB 1; Length 1466;
Best Local Similarity 27.8%; Pred. No. 4.6;
Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;

QY 5 PSDVSSCEDREGROLPAERPPQLRPGAPTSLQT--RP--QGND-----EGNHGGE 52
Db 973 PQGVKSGGKPGANGLSGERGPPGQGLAGTAGGPGDRGNPGSLGRGSGPGKG 1032
QY 53 DSCPHGSPQGPLAP--PASRGPATRSPLTFMRSSLLSRSSSGYFSFDTDRSPAPMSC 110
Db 1033 DRGENSGPAGAPGPHGPPCPV-----GPAKSG 1062

QY 111 DKSTQTFSPPCQA 123
Db 1063 DRGSGPAGPAGA 1075

RESULT 13
T26812

hypothetical protein Y41E3.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C;Accession: T26812
R;McMurray, A.
submitted to the EMBL Data Library, May 1997
A;Reference number: Z20270
A;Accession: T26812
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-289 <MIL>
A;Cross-references: EMBL:Z95559; PIDN:CAB08999.1; GSPDB:GN00022; CESP:Y41E3.2
A;Experimental source: clone Y41E3
C;Genetics:
A;Gene: CESP:Y41E3.2
A;Map position: 4
A;Introns: 27/3; 273/2
C;Superfamily: unassigned collagens

Query Match 9.2%; Score 97.5; DB 2; Length 289;
Best Local Similarity 35.4%; Pred. No. 1;
Matches 28; Conservative 6; Mismatches 34; Indels 11; Gaps 4;

QY 5 PSDVSSCEDREGROLPAE-----RPQO--LRPGAPTSLQTEPQG--NPEGNHGEGDSC 55
Db 97 PQGTPGKPGKPGKPGQPGTPTTTPPCQPCQPGPPGPGQPGIPGDNG 156
QY 56 PHG--SPQGGLAPPASPGP 72
Db 157 PPGGQPGKPGDAFGEPPG 175

RESULT 14
I49140
p62 ras-GAP associated phosphoprotein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I49140
R;Richard, S.; Yu, D.; Blumer, K.J.; Hausladen, D.; Olszowy, M.W.; Connelly, P.A.; Sh
Mol. Cell. Biol. 15, 186-197, 1995
A;Title: Association of p62, a multifunctional SH2- and SH3-domain-binding protein, w
A;Reference number: I49140; MUID:95097990
A;Accession: I49140
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-443 <RES>
A;Cross-references: EMBL:U17046; NID:g608527; PIDN:AAA64997.1; PID:g608528
C;Keywords: phosphoprotein

Query Match 9.2%; Score 97.5; DB 2; Length 443;
Best Local Similarity 27.0%; Pred. No. 1.7;
Matches 33; Conservative 20; Mismatches 54; Indels 15; Gaps 6;

QY 12 CDREGROLPAERPPQLRPGAPTSLQTEPQGNPEGNHGEGDSCPHGSPQGLAPPASPG 71
Db 19 CSDKPSGAHPSVR--LTPSRPSPLPHRPGGGGPRGARAS-PATQPP-PLLLPSTPG 73
QY 72 PFATRSPLTFMRSSLLSRSSSGYFSFDTD-RSPAPMSCDKSTQTPSPPCQAFNHVLSA 130
Db 74 PDAT-----VVGSAPTLLPFSATAAVKMEPNKYPELMAEKSLDPS-----FTHAMQL 124
QY 131 MA 132
Db 125 LS 126

RESULT 15
T10811
channel associated protein of synapse 2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000

C:Accession: T10811
 R:Irrie, M.; Hata, Y.; Takai, Y.
 submitted to the EMBL Data Library, April 1996
 A:Description: Cloning of new isoforms of PSD-95/SAP90 related genes.
 A:Reference number: Z17166
 A:Accession: T10811
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-852 <IRI>
 A:Cross-references: EMBL:U53368; NID:g1517939; PID:g1517940
 C:Genetics:
 A>Note: PSD-95/SAP90-related gene 2 (chapsyn isoform2)
 C:Superfamily: discs-large tumor suppressor; GLGF domain homology; guanylate kinase homology
 F:198-276/Domain: GLGF domain homology <GLG>
 F:543-601/Domain: SH3 homology <SH3>
 F:663-840/Domain: guanylate kinase homology <GKI>

Query Match 9.1%; Score 97; DB 2; Length 852;
 Best Local Similarity 23.8%; Pred. No. 3.7;
 Matches 44; Conservative 21; Mismatches 58; Indels 62; Gaps 9;

QY 28 LRPGAPTSL-QTEPGNPGNHGGEGDCSPH--GSPQGPL-----APPASPGPEA----- 74
 Db 274 LKVGRTTIYWDPIGPPDITHSYSPPMENHLLSGNNGTLEYKTSLLPISGRISPIPKH 333
 QY 75 -----TRSPLEIFMRRSSLLSRSSSYFSFTDRSPAPMSCDKS--TQTPSPPCQAFN 125
 Db 334 MLVEDEYTRPPEPVYSTVKNKCDKPAS-----PRHYSFVECDKSFLLSTYP----- 380
 QY 126 HYL-----SAMASRQAPADMRPEITWIAQELRIGDEFNAYARRVFLNNYQAAEDHP 179
 Db 381 HYHGLLPDSDMTSHSQHSTATROPSTVTLQRAI-----SLEGE 419
 QY 180 RMVIL 184
 Db 420 RKWL 424

Search completed: August 16, 2002, 10:01:47
 Job time: 140 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 10:01:22 ; Search time 11.99 Seconds
(without alignments)
639.405 Million cell updates/sec

Title: US-09-508-832-10
Perfect score: 1065
Sequence: 1 MAKOPSDVSECDREGRLQ.....PRMVILRLRYIVLVRMH 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	1065	100.0	198	1 BIM_HUMAN	O43521 homo sapien
2	908	85.3	196	1 BIM_MOUSE	O54918 mus musculus
3	893	83.8	196	1 BIM_RAT	O88498 rattus norv
4	111.5	10.5	886	1 CNA4_HUMAN	P27815 homo sapien
5	106.5	10.0	620	1 EXON_HSV2	P06489 herpes simp
6	102.5	9.6	1395	1 CUT1_MOUSE	P53564 mus musculus
7	100	9.4	627	1 ACHA_HUMAN	P43681 homo sapien
8	99	9.3	636	1 CA13_RAT	P13941 rattus norv
9	99	9.3	1464	1 CA13_MOUSE	P08121 mus musculus
10	99	9.3	1466	1 CA13_HUMAN	P02461 homo sapien
11	98	9.2	105	1 COLL_HSV7	P25050 herpesvirus
12	97.5	9.2	1505	1 CUT1_HUMAN	P39880 homo sapien
13	97	9.1	852	1 DLG2_RAT	O63622 rattus norv
14	96	9.0	870	1 DLG2_HUMAN	Q15700 homo sapien
15	96	9.0	1527	1 CA1H_MOUSE	P39061 mus musculus
16	95.5	9.0	210	1 CAC2_HAECO	P16252 haemonchus
17	95	8.9	298	1 CC34_CAEEL	P34687 caenorhabdi
18	95	8.9	844	1 CNA4_RAT	P54748 rattus norv
19	95	8.9	1516	1 CA1H_HUMAN	P39060 homo sapien
20	94.5	8.9	573	1 SECD_MYCTU	Q50634 mycobacteri
21	94.5	8.9	628	1 V70K_TYMW4	P20131 turnip yeil
22	94.5	8.9	1183	1 DRPL_RAT	P54258 rattus norv
23	94.5	8.9	1185	1 DRPL_HUMAN	P54259 homo sapien
24	94.5	8.9	1234	1 B3A2_RAT	P23347 rattus norv
25	94	8.8	580	1 SYN3_HUMAN	O14994 homo sapien
26	94	8.8	1460	1 CA11_CANFA	Q9xsj7 canis famill
27	93.5	8.8	902	1 NFCA_HUMAN	Q14934 homo sapien
28	92.5	8.7	245	1 C10A_HUMAN	P02745 homo sapien
29	92.5	8.7	408	1 AL_DROME	Q06453 drosophila
30	92.5	8.7	1446	1 IE18_PVKKA	P33479 pseudorabie
31	92.5	8.7	1461	1 IE18_PRVIF	P11675 pseudorabie
32	91.5	8.6	234	1 PRPM_HUMAN	P10161 homo sapien
33	91.5	8.6	276	1 PRPL_HUMAN	P10162 homo sapien

Query Match 100.0%; Score 1065; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 1e-68;

ALIGNMENTS

RESULT 1

ID	BIM_HUMAN	STANDARD;	PRT;	198 AA.
AC	O43521; O43522;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Bcl2-like protein 11 (Bcl2 interacting mediator of cell death).			
GN	BCL2L11 OR BIM.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid:9606;			
RP	[1]			
RN	SEQUENCE FROM N.A., FUNCTION, AND ALTERNATIVE SPLICING.			
RC	TISSUE=Peripheral blood, and Spleen;			
RX	MEDLINE=98094360; PubMed=9430630;			
RA	O'Connor L., Strasser A., O'Reilly L.A., Hausmann G., Adams J.M.,			
RA	Cory S., Huang D.C.S.;			
RT	*Bim: a novel member of the Bcl-2 family that promotes apoptosis.*;			
RL	EMBO J. 17:384-395(1998).			
CC	-!- FUNCTION: INDUCES APOPTOSIS. ISOFORM BIML IS MORE POTENT THAN			
CC	ISOFORM BIMEL.			
CC	-!- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2			
CC	PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES			
CC	NOT HETERODIMERIZE WITH PROAPOPTIC PROTEINS SUCH AS BAD, BOK,			
CC	BAX OR BAK (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOSOLASMIC MEMBRANES			
CC	(BY SIMILARITY).			
CC	-!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BIMEL (SHOWN HERE) AND			
CC	BIML; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND			
CC	CYTOTOXICITY.			
CC	-!- SIMILARITY: CONTAINS 1 BCL-2 HOMOLGY DOMAIN 3 (BH3).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AF032457; AAC39593.1; -.			
DR	EMBL; AF032458; AAC39594.1; -.			
DR	MIM; 603827; -.			
DR	InterPro; IPR000712; Bcl_2.			
DR	PROSITE; PS01259; BH3; FALSE_NEG.			
KW	Apoptosis; Alternative splicing; Membrane.			
FT	DOMAIN 148 162 BH3.			
FT	VARSPLIC 42 101 MISSING (IN ISOFORM BIML).			
SQ	SEQUENCE 198 AA; 22171 MW; D75735B469CA6997 CRC64;			

Q62245 mus musculus
P35247 homo sapien
P19484 homo sapien
P13808 mus musculus
O93353 gallus gall
P17600 homo sapien
P05997 homo sapien
P20186 streptomyce
P04258 bos taurus
Q02343 oryctolagus
Q05152 oryctolagus
O43365 homo sapien

34 91.5 8.6 1319 1 SOS1_MOUSE
35 91 8.5 375 1 TSPD_HUMAN
36 91 8.5 476 1 TFEH_HUMAN
37 91 8.5 1237 1 B3A2_MOUSE
38 90.5 8.5 413 1 HXD3_CHICK
39 90.5 8.5 705 1 SYN1_HUMAN
40 90.5 8.5 1496 1 CA25_HUMAN
41 90 8.5 348 1 YT35_STRFR
42 90 8.5 1049 1 CA13_BOVIN
43 90 8.5 2259 1 CCAE_RABIT
44 90 8.5 2339 1 CCAB_RABIT
45 89 8.4 443 1 HXA3_HUMAN

Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKQPSVSSCEDREGQLQPAERPPQLRPGAPTSLOTQPOGNEGHNHGGEGDSCPHGSP 60
 DB 1 MAKQPSVSSCEDREGQLQPAERPPQLRPGAPTSLOTQPOGNEGHNHGGEGDSCPHGSP 60

QY 61 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120
 DB 61 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120

QY 121 COAFNHVLSAMASMRQA--EPADMREIWIQAELRRIGDEFNAYYARRVFLNYYQAEDHPR 180
 DB 121 COAFNHVLSAMASMRQA--EPADMREIWIQAELRRIGDEFNAYYARRVFLNYYQAEDHPR 180

QY 181 MVILRLRYIVRLVWRMH 198
 DB 181 MVILRLRYIVRLVWRMH 198

RESULT 2

BIM_MOUSE
 ID BIM_MOUSE STANDARD; PRT; 196 AA.
 AC 054918; 054919; 054920;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE BCL2-like protein 11 (BCL2 interacting mediator of cell death).
 GN BCL2L11 OR BIM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE
 SPECIFICITY, AND ALTERNATIVE SPLICING.
 RX MEDLINE=98094360; PubMed=9430630;
 RA O'Connor L., Strasser A., O'Reilly L.A., Hausmann G., Adams J.M.,
 RA Cory S., Huang D.C.S.;
 RA "Bim: a novel member of the Bcl-2 family that promotes apoptosis.";
 RL EMBO J. 17:384-395(1998).
 CC -1- FUNCTION: INDUCES APOPTOSIS. THE ISOFORMS VARY IN CYTOTOXICITY
 WITH ISOFORM BIMs BEING THE MOST POTENT AND ISOFORM BIMEL BEING
 THE LEAST POTENT.
 CC -1- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2
 CC PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES
 CC NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,
 CC BAX OR BAK (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOSOLIC MEMBRANES.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; BIMEL (SHOWN HERE), BIML AND
 CC BIMs; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF B- AND T-LYMPHOID CELL
 CC LINES.
 CC -1- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND
 CC CYTOTOXICITY.
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLGY DOMAIN 3 (BH3).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF032459; AAC40029.1; -
 CC EMBL: AF032460; AAC40030.1; -
 CC EMBL: AF032461; AAC40031.1; -
 CC MGD; MGI:1197519; Bcl2l11.
 CC InterPro: IPR000712; Bcl_2.
 CC ProSITE; PS01259; BH3; FALSE_NEG.
 CC Apoptosis; Alternative splicing; Membrane.
 FT DOMAIN 146 160 BH3
 FT VARSPPLIC 42 97 MISSING (IN ISOFORM BIML).

FT VARSPPLIC 42 127 MISSING (IN ISOFORM BIMs).
 SQ SEQUENCE 196 AA; 22066 MW; 531C176E5FIAC9AA CRC64;

Query Match 85.3%; Score 908; DB 1; Length 196;
 Best Local Similarity 86.5%; Pred. No. 1.1e-57;
 Matches 173; Conservative 10; Mismatches 11; Indels 6; Gaps 2;

QY 1 MAKQPSVSSCEDREGQLQPAERPPQLRPGAPTSLOTQPOGNEGHNHGGEGDSCPHGSP 60
 DB 1 MAKQPSVSSCEDREGQLQPAERPPQLRPGAPTSLOTQPOGNEGHNHGGEGDSCPHGSP 56

QY 61 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 120
 DB 57 QGLAPPASPGPFATRSPLFTFMRRSSLLSSSSSGYFSFDTDRSPAPMSCDKSTQTPSP 116

QY 121 COAFNHVLSAMASMRQA--EPADMREIWIQAELRRIGDEFNAYYARRVFLNYYQAEDH 178
 DB 117 COAFNHVLSAMASIRQSEEPEDLRPEIRIAQELRRIGDEFNYYTRRVFANDYREAEDH 176

QY 179 PRVILRLRYIVRLVWRMH 198
 DB 177 PQWVILQLLRIFELVWRH 196

RESULT 3

BIM_RAT
 ID BIM_RAT STANDARD; PRT; 196 AA.
 AC 088498; 09WU18; 088497;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE BCL2-like protein 11 (BCL2 interacting mediator of cell death)
 DE (Bcl-2 related ovarian death protein).
 GN BCL2L11 OR BIM OR BOD.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND TISSUE SPECIFICITY
 RP (ISOFORMS BOD-L; BOD-M AND BOD-S).
 RC TISSUE=Ovary;
 RX MEDLINE=98400436; PubMed=9731710;
 RA Hsu S.Y., Lin P., Hsueh A.J.W.;
 RA "BOD (Bcl-2-related ovarian death gene) is an ovarian BH3 domain-
 RT containing proapoptotic Bcl-2 protein capable of dimerization with
 RT diverse antiapoptotic Bcl-2 members.";
 RL Mol. Endocrinol. 12:1432-1440(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM BIML).
 RP Chen D., Simon R.P., Chen J.;
 RT "Cloning of rat bimEL and bimL, and their differential expression in
 RT ischemia and normal rat brain.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INDUCES APOPTOSIS.
 CC -1- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2
 CC PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES
 CC NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,
 CC BAX OR BAK.
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOSOLIC MEMBRANES
 CC (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; BOD-L (SHOWN HERE) AND BOD-S;
 CC ARE PRODUCED BY THE USE OF ALTERNATIVE INITIATION SITES. TWO
 CC FURTHER ISOFORMS; BIML AND BOD-M; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING OF BOD-L.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND
 CC CYTOTOXICITY.
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLGY DOMAIN 3 (BH3).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF032459; AAC40029.1; -
 CC EMBL: AF032460; AAC40030.1; -
 CC EMBL: AF032461; AAC40031.1; -
 CC MGD; MGI:1197519; Bcl2l11.
 CC InterPro: IPR000712; Bcl_2.
 CC ProSITE; PS01259; BH3; FALSE_NEG.
 CC Apoptosis; Alternative splicing; Membrane.
 FT DOMAIN 146 160 BH3
 FT VARSPPLIC 42 97 MISSING (IN ISOFORM BIML).

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF065433; AAC23595.1; -
DR EMBL; AF065431; AAC23593.1; -
DR EMBL; AF065432; AAC23594.1; -
DR EMBL; AF136927; AAD26594.1; -
DR InterPro: IPR000712; BCL_2
DR PROSITE; PS01259; BH3; FALSE_NEG.
KW Apoptosis; Alternative splicing; Membrane; Alternative initiation.
FT CHAIN 1 196 BCL2-LIKE PROTEIN 11, ISOFORM BOD-L.
FT CHAIN 104 196 BCL2-LIKE PROTEIN 11, ISOFORM BOD-S.
FT INIT_MET 104 104 FOR ISOFORM BOD-S.
FT DOMAIN 146 160 BH3.
FT VARSPIC 42 97 MISSING (IN ISOFORM B1ML).
FT VARSPIC 42 137 MISSING (IN ISOFORM BOD-M).
FT CONFLICT 136 136 E -> D (IN REF. 1; AAC23594).
SQ SEQUENCE 196 AA; 22055 MW; B4D2146F9C0B37A0 CRC64;

Query Match 83.8%; Score 893; DB 1; Length 196;
Best Local Similarity 85.0%; Pred. No. 1.2e-56;
Matches 170; Conservative 11; Mismatches 13; Indels 6; Gaps 2;

QY 1 MAKOPSDVSECDREGCOLQPAERPPQLRPGAPTSLOTPEQNGHGGDCSPHGSP 60
Db 1 MAKOPSDVSECDREGCOLQPAERPPQLRPGAPTSLOTPEQNGHGGDCSPHGSP 56
QY 61 QGELAPPASPGPATRSPFLIFMRRSLLSRSSGYFSDTDRSPAPMSCDKSTQTPSP 120
Db 57 QGELAPPASPGPATRSPFLIFVRRSLLSRSSGYFSDTDRSPAPMSCDKSTQTPSP 116
QY 121 QCAFNYLSMASMROA--EPADMRREITWQELRIGRIGDEFNAYYARRVPLNNYQAEHD 178
Db 117 QCAFNYLSMASIRSQEPEDLRPEIRTAQELRIGRIGDEFNAYYARRVPLNNYQAEHD 176
QY 179 PRVILRLRLRYIVRLVWRMH 198
Db 177 PQWVILQLLRFIFRLVWRH 196

RESULT 4
CN4A_HUMAN STANDARD: PRT: 886 AA.
ID CN4A_HUMAN Q16255; O75222; O76092;
AC P27815; Q16691; Q16255; O75222; O76092;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CAMP-dependent 3',5'-cyclic phosphodiesterase 4A (EC 3.1.4.17)
DE (DPDE2) (PDE46).
GN PDE4A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94019330; PubMed=8413254;
RA Bolger G., Michaeli T., Martins T., St John T., Steiner B.,
RA Rodgers L., Riggs M., Wigler M., Ferguson K.;
RT "A family of human phosphodiesterases homologous to the dunce
RT learning and memory gene product of *Drosophila melanogaster* are
RT potential targets for antidepressant drugs.";
RL Mol. Cell. Biol. 13:6558-6571(1993).
RN [2]
RP SEQUENCE FROM N.A. (PDE4A RD1).
RC TISSUE=Brain;
RX MEDLINE=98343959; PubMed=9677330;
RA Sullivan M., Rana G., Begg F., Gordon L., Olsen A.S., Houslay M.D.;
RT "Identification and characterization of the human homologue of the

RT short PDE4A CAMP-specific phosphodiesterase RD1 (PDE4A1) by analysis
RT of the human HSPD54A gene locus located at chromosome 19p13.2.";
RL Biochem. J. 333:693-703(1998).
RN [3]
RP SEQUENCE OF 112-886 FROM N.A.
RC TISSUE=Monocytes;
RX MEDLINE=90258854; PubMed=2160582;
RA Livi G.P., Kmetz P., McHale M.M., Cieslinski L.B., Sathe G.M.,
RA Taylor D.P., Davis R.L., Torphy T.J., Balcarek J.M.;
RT "Cloning and expression of cDNA for a human low-Km, rolipram-sensitive
RT cyclic AMP phosphodiesterase.";
RL Mol. Cell. Biol. 10:2678-2686(1990).
RN [4]
RP REVISIONS TO REF.3.
RA McLaughlin M.M.;
RL Submitted (JAN-1991) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (PDE4A7/A6 AND PDE4A8/2EL).
RX MEDLINE=95290008; PubMed=7772058;
RA Hordton Y.M., Sullivan M., Houslay M.D.;
RT "Molecular cloning of a novel splice variant of human type IVA
RT (PDE-IVA) cyclic AMP phosphodiesterase and localization of the gene
RT to the p13.2-q12 region of human chromosome 19.";
RL Biochem. J. 308:683-691(1995).
RN [6]
RP SEQUENCE FROM N.A. (PDE4A7/A6).
RX MEDLINE=95194817; PubMed=7888306;
RA Sullivan M., Egerton M., Shakur Y., Marquardsen A., Houslay M.D.;
RT "Molecular cloning and expression, in both COS-1 cells and S.
RT cerevisiae, of a human cytosolic type-IVA, cyclic AMP specific
RT phosphodiesterase (hPDE-IVA-h6.1).";
RL Cell. Signal. 6:793-812(1994).
CC -!- CATALYTIC ACTIVITY: ADENOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O -
CC ADENOSINE 5'-PHOSPHATE.
CC -!- ENZYME REGULATION: INHIBITED BY ROLIPRAM.
CC -!- PATHWAY: CYCLIC NUCLEOTIDE METABOLISM.
CC -!- SUBCELLULAR LOCATION: THE PDE4A1 RD1 ISOFORM HAS PROPENSITY FOR
CC ASSOCIATION WITH MEMBRANES.
CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS OF THE PROTEIN ARE PRODUCED BY
CC AND PDE4A1/RD1 A SHORT ISOFORM. ISOFORM 2EL PROBABLY REPRESENTS A
CC NON-FUNCTIONAL SPICE VARIANT. THE PRESENCE OF DIFFERENT N-
CC TERMINAL REGIONS HAS BEEN LINKED WITH DISTINCT FUNCTIONAL ROLES.
CC -!- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L20965; AAA03588.1; -
DR EMBL; AF069491; AAC35012.1; -
DR EMBL; AF069487; AAC35012.1; JOINED.
DR EMBL; AF069489; AAC35012.1; JOINED.
DR EMBL; AF069490; AAC35012.1; JOINED.
DR EMBL; AF069491; AAC35013.1; -
DR EMBL; AF069489; AAC35013.1; JOINED.
DR EMBL; AF069490; AAC35013.1; JOINED.
DR EMBL; AF069491; AAC35014.1; -
DR EMBL; AF069489; AAC35014.1; JOINED.
DR EMBL; AF069490; AAC35014.1; JOINED.
DR EMBL; M37744; AAA69697.1; -
DR EMBL; U18087; AAC50458.1; -
DR EMBL; U18088; AAC98540.1; -
DR EMBL; S75213; AAB33798.1; -
DR EMBL; AF069491; AAC35015.1; -
DR EMBL; AF069488; AAC35015.1; JOINED.
DR EMBL; AF069489; AAC35015.1; JOINED.
DR EMBL; AF069490; AAC35015.1; JOINED.

[2]
SEQUENCE OF 64-1395 FROM N.A. (ISOFORM 1).
RC STRAIN-A/J. AND BALB/C: TISSUE=Brain;
RX MEDLINE=94244481; PubMed=7910552;
RA Valarche I.; Tissier-Seca J.P.; Hirsch M.R.; Martinez S.; Goridis C.,
Brunet J.F.;
RT "The mouse homeodomain protein Phox2 regulates Ncam promoter activity
in concert with Cux/CDP and is a putative determinant of
neurotransmitter phenotype.";
RL Development 119:881-896(1993).
RN [3]
SEQUENCE OF 642-1395 FROM N.A.
RX MEDLINE=96437626; PubMed=8840273;
RA Den Heuvel G.B., Bodmer R., McConnell K.R., Nagami G.T., Igarashi P.;
RT "Expression of a cut-related homeobox gene in developing and
polycystic mouse kidney.";
RL Kidney Int. 50:453-461(1996).
RN [4]
SEQUENCE OF 936-1395 FROM N.A.
RC TISSUE=Testis;
RA Quaglin S.E., Igarashi P.;
RT "A unique variant of a homeobox gene related to Drosophila cut is
expressed in mouse testis.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A
REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY
PREVENTING BINDING OF POSITIVELY-ACTING CCAAT FACTORS TO
PROMOTERS (BY SIMILARITY). COMPONENT OF NF-MUNR REPRESSOR; BINDS
TO THE MARS (5' AND 3') OF THE IMMUGLOBULIN HEAVY CHAIN
ENHANCER.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AF004225; AAD12485.1; -
DR EMBL; X75013; CAA52922.1; -
DR EMBL; U46683; AAC52775.1; -
DR EMBL; U46684; AAB41146.1; -
DR HSSP; P10037; IAU7.
DR MGD; MGI:88568; Cut11.
DR InterPro; IPR003350; CUT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF02376; CUT; 3.
DR Pfam; PF00046; homeobox; 2.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Transcription regulation; Homeobox; DNA-binding;
KW Developmental protein; Nuclear protein; Repeat; Repressor;
KW Coiled coil; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 243 COILED COIL (POTENTIAL).
FT DNA_BIND 420 507 CUT 1.
FT DOMAIN 547 603 COILED COIL (POTENTIAL).
FT DNA_BIND 809 896 CUT 2.
FT DNA_BIND 992 1079 CUT 3.
FT DNA_BIND 1119 1178 HOMEBOX.
FT VARSPPLIC 287 388 MISSING (IN ISOFORM 2).
FT CONFLICT 1360 1360 G -> A (IN REF. 2).
FT CONFLICT 1385 1385 P -> L (IN REF. 1).
SQ SEQUENCE 1395 AA; 151802 MW; D062CC227D7A163E CRC64;

Query Match 9.6%; Score 102.5; DB 1; Length 1395;
Best Local Similarity 23.7%; Pred. No. 3.5;
Matches 54; Conservative 25; Mismatches 76; Indels 73; Gaps 9;
QY 6 SDVSECDREGHQLOPAERPPOLRPGAPTSLQT---EPQGNPEGNHGEGDSCPHGSPG 62
Db 282 SDLSGSARKGRDQESRRPGLPASPPLPRNTGEQVSTNGTH-----HFSPAG 333
QY 63 -----PLAPPASPGPFATRSPLFIEMRS-----SLLSRSSSGVFSFD 100
Db 334 LSQDFSSNLASPSLPLASTGRFALNLLQRLQMOSFYSKAMQAEAGSTSTIFSTGYPSTN 393
QY 101 TDRSPAPM---SCDKSTQTPTSP--- PRT; 627 AA.
Db 394 SISSPSLQOSPDVNCMAPSPQSSESAGSISGEEIDTAETARQVKEQLIKHINIGORIFG 453
QY 126 HYLSSAMASMRQAEPADMRPEIWIQAELRIGDEFNAYARRVFLNNYQ 173
Db 454 HYLGLSQGSVSEIL--ARPKWNKLTVR--GKE--PFHKMKOFLSDEQ 496
RESULT 7
ACH4_HUMAN
ID AC4_HUMAN STANDARD; PRT; 627 AA.
AC P43681;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-4 chain precursor.
GN CHRNA4 OR NACRA4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95237608; PubMed=7721089;
RA Monteggia L.M., Gopalakrishnan M., Touma E., Idler K.B., Nash N.,
Arneric S.P., Sullivan J.P., Giordano T.;
RT "Cloning and transient expression of genes encoding the human alpha-4
and beta-2 neuronal nicotinic acetylcholine receptor (nAChR)
subunits.";
RT Gene 155:189-193(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96430009; PubMed=88333159;
RA Steinlein O.K., Weiland S., Stood J., Propping P.;
RT "Exon-intron structure of the human neuronal nicotinic acetylcholine
receptor alpha 4 subunit (CHRNA4).";
RL Genomics 32:289-294(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=97062879; PubMed=8906617;
RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,
Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;
RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
2-beta 4 nicotinic acetylcholine receptor subunits and functional
expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
beta 4 subunits.";
RL J. Mol. Neurosci. 7:217-228(1996).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97162233; PubMed=9009220;
RA Groot Kormelink P.J., Luyten W.H.M.L.;
RT "Cloning and sequence of full-length cDNAs encoding the human neuronal
nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and
expression of seven nAChR subunits in the human neuroblastoma cell
line SH-SY5Y and/or IMR-32.";
RL FEBS Lett. 400:309-314(1997).
RN [5]
RP SEQUENCE FROM N.A.

[illegible]

```

RA Komm B., Mohn K.;
RT "Regulation of alpha 2(I), alpha 1(III), and alpha 2(V) collagen
RL mRNAs by estradiol in the immature rat uterus.";
CC DNA 7:347-354(1988).
CC -!- FUNCTION: COLLAGEN.
CC -!- ALONG WITH TYPE I COLLAGEN.
CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70369; CAA49832.1; -.
DR EMBL; AJ005395; CAA06510.1; -.
DR EMBL; M21354; AAA40942.1; -.
DR PIR; A29905; A29905.
DR PIR; S41067; S41067.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib.collagen_C.
DR InterPro; IPR001007; VWFC.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 6.
DR ProDom; PD002078; Fib.collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR PROSITE; PS01208; VWFC; PARTIAL.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Collagen; Glycoprotein.
FT NON_TER 1
FT CHAIN <1 375 COLLAGEN ALPHA 1(III) CHAIN.
FT PROPEP 376 636 CARBOXYL-TERMINAL PROPEPTIDE.
FT DOMAIN <1 368 TRIPLE-HELICAL REGION.
FT DOMAIN 369 636 NONHELICAL REGION (C-TERMINAL).
FT DISULFID 368 368 INTERCHAIN (BY SIMILARITY).
FT DISULFID 369 369 INTERCHAIN (BY SIMILARITY).
FT CONFLICT 340 340 N -> D (IN REF. 2).
FT CONFLICT 429 429 A -> G (IN REF. 2).
SQ SEQUENCE 636 AA; 62332 MW; 61AA48159F01D01EE CRC64;

Query Match 9.3%; Score 99; DB 1; Length 636;
Best Local Similarity 35.4%; Pred. No. 2.9;
Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;

QY 5 PSDVSSECDREGROLQPAER---PPQLRPGAP-TSLQTEPGGNP-----EGNHGGEG 52
D 145 PQGKESKPGKNGSNGRGGPGGLPGQPGTAGPGRDGNPGSDGPGRDSGPGGK 204
QY 53 DSCPFGSPQGPLAP--PASPGP 72
D 205 DRNGSPGAPGAPGHPGPPG 226

RESULT 9
CA13_MOUSE
ID CA13_MOUSE STANDARD; PRT; 1464 AA.
AC P08121; O61429; O9CRN7;
DT 01-AUG-1988 (Rel. 08, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 1(III) chain precursor.
GN COL3A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL X DBA; TISSUE=Embryo;
RX MEDLINE=95011609; PubMed=7926795;
RA Toman B., de Crombrughe B.;
RT "The mouse type-III procollagen-encoding gene: genomic cloning and
RL complete DNA sequence.";
RL Gene 147:161-168(1994).
RN [2]
RP SEQUENCE OF 1-488 FROM N.A.
RX MEDLINE=88167858; PubMed=3443309;
RA Wood L., Theriault N., Vogel G.;
RT "Complete nucleotide sequence of the N-terminal domains of the murine
RL alpha-1 type-III collagen chain.";
RL Gene 61:225-230(1987).
RN [3]
RP SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=85131189; PubMed=3972847;
RA Liau G., Mudryj M., de Crombrughe B.;
RT "Identification of the promoter and first exon of the mouse alpha 1
RL (III) collagen gene.";
RL J. Biol. Chem. 260:3773-3777(1985).
RN [4]
RP SEQUENCE OF 810-1464 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Yashaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP SEQUENCE OF 1442-1464 FROM N.A.
RC STRAIN=C57BL;
RX MEDLINE=91274355; PubMed=2054384;
RA Metsaeranta M., Toman D., de Crombrughe B., Vuorio E.;
RT "Specific hybridization probes for mouse type I, II, III and IX
RL collagen mRNAs.";
CC -!- FUNCTION: COLLAGEN TYPE III OCCURS IN MOST SOFT CONNECTIVE TISSUES
CC ALONG WITH TYPE I COLLAGEN.
CC -!- SUBUNIT: TRIMERS OF IDENTICAL ALPHA 1(III) CHAINS. THE CHAINS ARE
CC LINKED TO EACH OTHER BY INTERCHAIN DISULFIDE BONDS. TRIMERS ARE
CC ALSO CROSS-LINKED VIA HYDROXYLYSINES.
CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC O-LINKED GLYCAN CONSISTS OF GLC-GAL DISACCHARIDE (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52046; CAA36279.1; -.

```

```
DR EMBL; M18933; AAA37338.1; -.
DR EMBL; K03037; -, NOT_ANNOTATED_CDS.
DR EMBL; AK019448; BAB31724.1; -.
DR EMBL; X57983; CAA41048.1; -.
DR PIR; A22287; A22287.
DR PIR; A27353; A27353.
DR PIR; S16373; S16373.
DR MGD; MGI:88453; Col3a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000885; Fib.collagen_C.
DR InterPro; IPR001007; WFEC.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 17.
DR ProDom; PD002078; Fib.collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS01208; WVC; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Collagen; Signal.
FT SIGNAL 1 23 BY SIMILARITY.
FT PROPEP 24 154 AMINO-TERMINAL PROPEPTIDE.
FT CHAIN 155 1203 COLLAGEN ALPHA 1(III) CHAIN.
FT PROPEP 1204 1464 CARBOXYL-TERMINAL PROPEPTIDE.
FT DOMAIN 31 90 WVEC.
FT DOMAIN 155 169 NONHELICAL REGION (N-TERMINAL).
FT DOMAIN 170 1195 TRIPLE-HELICAL REGION.
FT DOMAIN 1196 1464 NONHELICAL REGION (C-TERMINAL).
FT CARBOHYD 262 262 O-LINKED (GAL. . .) (BY SIMILARITY).
FT MOD_RES 262 262 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 283 283 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 859 859 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 976 976 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 1093 1093 HYDROXYLATION (BY SIMILARITY).
FT MOD_RES 1105 1105 HYDROXYLATION (BY SIMILARITY).
FT DISULFID 1195 1195 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1196 1196 INTERCHAIN (BY SIMILARITY).
FT SEQUENCE 1464 AA; 138944 MW; 2104EC27A886090B CRC64;

Query Match          9.3%; Score 99; DB 1; Length 1464;
Best Local Similarity 35.4%; Pred. No. 6.5;
Matches 29; Conservative 8; Mismatches 31; Indels 14; Gaps 4;

QY 5 PSDVSSCEDREGROLQPAER---PQLRPGAP-TSLTEPQGNP-----EGNHGEG 52
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 972 PQGFKGESGKPGASGNGERPPGQGLPGQPGTAGEPGRDGNPGSDGQPGRDGSPGKG 1031
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 53 DSCPHGSPQGPLAP--PASPCP 72
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1032 DRGENSGPAGAPGHPGPPGP 1053

RESULT 10
CAL3_HUMAN
ID CAL3_HUMAN STANDARD; PRT; 1466 AA.
AC P02461; Q15112;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 1(III) chain precursor.
GN COL3A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin fibroblast;
RX MEDLINE=89350838; PubMed=2764886;
RA Ala-Kokko L., Kontusaari S., Baldwin C.T., Kuivaniemi H.,
RA Prockop D.J.;
RT "Structure of cDNA clones coding for the entire prepro alpha 1 (III)
   chain of human type III procollagen. Differences in protein structure
```

```
RT from type I procollagen and conservation of codon preferences.";
RL Biochem. J. 260:509-516(1989).
RX [2]
RX SEQUENCE OF 149-1225 FROM N.A.
RX MEDLINE=89386015; PubMed=2780304;
RA Janeczko R.A., Ramirez F.;
RT "Nucleotide and amino acid sequences of the entire human alpha 1
   (III) collagen.";
RL Nucleic Acids Res. 17:6742-6742(1989).
RX [3]
RX SEQUENCE OF 168-398.
RX MEDLINE=77134724; PubMed=557335;
RA Seyer J.M., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of cyanogen
   bromide peptides from the amino-terminal segment of type III collagen
   of human liver.";
RL Biochemistry 16:1158-1164(1977).
RX [4]
RX REVISIONS.
RA Seyer J.M.;
RL Submitted (DEC-1977) to the PIR data bank.
RX [5]
RX SEQUENCE OF 399-727.
RX MEDLINE=79000343; PubMed=687591;
RA Seyer J.M., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of five
   consecutive CNBr peptides from type III collagen of human liver.";
RL Biochemistry 17:3404-3411(1978).
RX [6]
RX SEQUENCE OF 728-964.
RX MEDLINE=80198282; PubMed=6246925;
RA Seyer J.M., Mainardi C., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of alpha 1
   (III)-CB5 from type III collagen of human liver.";
RL Biochemistry 19:1583-1589(1980).
RX [7]
RX SEQUENCE OF 950-1466 FROM N.A.
RX MEDLINE=88189827; PubMed=3357782;
RA Mankoo B.S., Dalglish R.;
RT "Human pro alpha 1(III) collagen: cDNA sequence for the 3' end.";
RL Nucleic Acids Res. 16:2337-2337(1988).
RX [8]
RX REVISION TO 1184.
RX MEDLINE=89098346; PubMed=3211760;
RA Molyneux K., Dalglish R.;
RT "Human type III collagen 'variant' is a cDNA cloning artefact.";
RL Nucleic Acids Res. 16:11833-11833(1988).
RX [9]
RX SEQUENCE OF 1065-1466 FROM N.A.
RX MEDLINE=85087944; PubMed=6096827;
RA Loidl H.R., Brinker J.M., May M., Pihlajaniemi T., Morrow S.,
RA Rosenbloom J., Myers J.C.;
RT "Molecular cloning and carboxyl-propeptide analysis of human type III
   procollagen.";
RL Nucleic Acids Res. 12:9383-9394(1984).
RX [10]
RX SEQUENCE OF 965-1200.
RX MEDLINE=81208139; PubMed=7016180;
RA Seyer J.M., Kang A.H.;
RT "Covalent structure of collagen: amino acid sequence of alpha
   1(III)-CB9 from type III collagen of human liver.";
RL Biochemistry 20:2621-2627(1981).
RX [11]
RX SEQUENCE OF 1176-1466 FROM N.A.
RX MEDLINE=85157600; PubMed=2579949;
RA Chu M.-L., Weil D., de Wet W.J., Bernard M.P., Sippola M., Ramirez F.;
RT "Isolation of cDNA and genomic clones encoding human pro-alpha 1
   (III) collagen. Partial characterization of the 3' end region of the
   gene.";
RL J. Biol. Chem. 260:4357-4363(1985).
RX [12]
RX SEQUENCE OF 1161-1200 FROM N.A.
RX MEDLINE=86187804; PubMed=3734462;
```

RA Miskulin M., Dalgleish R., Kluve-Beckerman B., Rennard S.I.,
 RA Tolstoshev P., Brantly M., Crystal R.G.;
 RA "Human type III collagen gene expression is coordinately modulated
 RT with the type I collagen genes during fibroblast growth.";
 RL Biochemistry 25:1408-1413(1986).
 RN [13]
 RP SEQUENCE OF 1-170 FROM N.A.
 RC TISSUE-Placenta;
 RX MEDLINE=88303360; PubMed=3405773;
 RA Tomán D., Ricca G., de Crombrughe B.;
 RT "Nucleotide sequence of a cDNA coding for the amino-terminal region
 RT of human prepro alpha 1(III) collagen.";
 RL Nucleic Acids Res. 16:7201-7201(1988).
 RN [14]
 RP SEQUENCE OF 1-176 FROM N.A.
 RX MEDLINE=89378752; PubMed=2777083;
 RA Benson-Chanda V., Su M.W., Weill D., Chu M.-L., Ramirez F.;
 RT "Cloning and analysis of the 5' portion of the human type-III
 RT procollagen gene (COL3A1).";
 RL Gene 78:255-265(1989).
 RN [15]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=97255959; PubMed=9101290;
 RA Kuivaniemi H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 RT associated collagen (type IX), and network-forming collagen (type X)
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315(1997).
 RN [16]
 RP VARIANT AORTIC ANEURYSM ARG-303, AND VARIANT THR-668.
 RX MEDLINE=93293988; PubMed=8514866;
 RA Tromp G., Wu Y., Prockop D.J., Madhathari S.L., Kleinert C.,
 RA Earley J.J., Zhuang J., Noerregaard O., Darling R.C., Abbott W.M.,
 RA Cole C.W., Jaakkola P., Rynnaenen M., Pearce W.H., Yao J.S.T., S.A.,
 RA Majamaa K., Smullens S.N., Gatalica Z., Ferrell R.E., Jimenez S.A.,
 RA Jackson C.E., Michels V.V., Kaye M., Kuivaniemi H.;
 RT "Sequencing of cDNA from 50 unrelated patients reveals that mutations
 RT in the triple-helical domain of type III procollagen are an
 RT infrequent cause of aortic aneurysms.";
 RL J. Clin. Invest. 91:2539-2545(1993).
 RN [17]
 RP VARIANT THR-698.
 RX MEDLINE=91045136; PubMed=2235526;
 RA Zafarullah K., Kleinert C., Tromp G., Kuivaniemi H., Kontusaari S.,
 RA Wu Y., Ganguly A., Prockop D.J.;
 RT "G to A polymorphism in exon 31 of the COL3A1 gene.";
 RL Nucleic Acids Res. 18:6180-6180(1990).
 RN [18]
 RP VARIANT AORTIC ANEURYSM ARG-786.
 RX MEDLINE=91056145; PubMed=2243125;
 RA Kontusaari S., Tromp G., Kuivaniemi H., Romanic A.M., Prockop D.J.;
 RT "A mutation in the gene for type III procollagen (COL3A1) in a family
 RT with aortic aneurysms.";
 RL J. Clin. Invest. 86:1465-1473(1990).
 RN [19]
 RP VARIANT EDS-IV ARG-828.
 RX MEDLINE=94016385; PubMed=8411057;
 RA Richards A.J., Narcisi P., Lloyd J.C., Ferguson C., Pope F.M.;
 RT "The substitution of glycine 661 by arginine in type III collagen
 RT produces mutant molecules with different thermal stabilities and
 RT causes Ehlers-Danlos syndrome type IV.";
 RL J. Med. Genet. 30:690-693(1993).
 RN [20]
 RP VARIANT EDS-IV SER-957.
 RX MEDLINE=89109135; PubMed=2492273;
 RA Tromp G., Kuivaniemi H., Shikata H., Prockop D.J.;
 RT "A single base mutation that substitutes serine for glycine 790 of
 RT the alpha 1 (III) chain of type III procollagen exposes an arginine
 RT and causes Ehlers-Danlos syndrome IV.";
 RL J. Biol. Chem. 264:1349-1352(1989).
 RN [21]
 RP VARIANT EDS-IV VAL-960.
 RX MEDLINE=95268429; PubMed=7749417;

RA Tromp G., de Paeppe A., Nuytinck L., Madhathari S.L., Kuivaniemi H.;
 RT "Substitution of valine for glycine 793 in type III procollagen in
 RT Ehlers-Danlos syndrome type IV.";
 RL Hum. Mutat. 5:179-181(1995).
 RN [22]
 RP VARIANT EDS-IV GLU-1014.
 RX MEDLINE=92316511; PubMed=1352273;
 RA Richards A.J., Ward P.N., Narcisi P., Nicholls A.C., Lloyd J.C.,
 RA Pope F.M.;
 RT "A single base mutation in the gene for type III collagen (COL3A1)
 RT converts glycine 847 to glutamic acid in a family with Ehlers-Danlos
 RT syndrome type IV. An unaffected family member is mosaic for the
 RT mutation.";
 RL Hum. Genet. 89:414-418(1992).
 RN [23]
 RP VARIANT EDS-IV ASP-1050.
 RX MEDLINE=90037070; PubMed=2808425;
 RA Tromp G., Kuivaniemi H., Stolle C.A., Pope F.M., Prockop D.J.;
 RT "Single base mutation in the type III procollagen gene that converts
 RT the codon for glycine 883 to aspartate in a mild variant of
 RT Ehlers-Danlos syndrome IV.";
 RL J. Biol. Chem. 264:19313-19317(1989).
 RN [24]
 RP VARIANT EDS-IV VAL-1077.
 RX MEDLINE=91374480; PubMed=1895316;
 RA Richards A.J., Lloyd J.C., Ward P.N., de Paeppe A., Narcisi P.,
 RA Pope F.M.;
 RT "Characterisation of a glycine to valine substitution at amino acid
 RT position 910 of the triple helical region of type III collagen in a
 RT patient with Ehlers-Danlos syndrome type IV.";
 RL J. Med. Genet. 28:458-463(1991).
 RN [25]
 RP VARIANT EDS-IV GLU-1173.
 RX MEDLINE=93022543; PubMed=1357232;
 RA Johnson P.H., Richards A.J., Pope F.M., Hopkinson D.A.;
 RT "Query Match 9.3%; Score 99; DB 1; Length 1466;
 RT Best Local Similarity 27.8%; Pred. No. 6.5;
 RT Matches 37; Conservative 7; Mismatches 45; Indels 44; Gaps 5;
 QY 5 PSDVSSECDREGRLQPAERPPQLRPGAPTSLQT--EP--QGNP-----EGNHGGEG 52
 DB 973 PGVKGESGKPGANGSLGSGRPPGPGQGLPGLAGTAGEDGPNPGDGLPGRDGSPPGKG 1032
 QY 53 DSCPHGSPGGLAP--PASPGFATRSPLFIFMRSSLLSRSSSGYFSFDTDRSPAPMSC 110
 DB 1033 DRGENSGPAGAPGHPGPGPV-----GPGAGKSG 1062
 QY 111 DXSTQTPSPPCQA 123
 DB 1063 DRGEGSPAGPAGA 1075
 RESULT 11
 COLL_HSVS7
 ID COLL_HSVS7 STANDARD; PRT; 105 AA.
 AC P25050;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Collagen-like protein.
 OS Herpesvirus saimiri (strain 484-77).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=10382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90279084; PubMed=2161952;
 RA Geck P., Whitaker S.A., Medveczky M.M., Medveczky P.G.;
 RT "Expression of collagenlike sequences by a tumor virus, herpesvirus
 RT saimiri.";
 RL J. Virol. 64:3509-3515(1990).
 RN [2]

```

RP REVISIONS.
RX MEDLINE=92046383; PubMed=1658399;
RA Geck P., Whitaker S.A., Medveczky M.M., Medveczky P.G.;
RT "Expression of collagenlike sequences by a tumor virus, herpesvirus
RT saimiri";
RL J. Virol. 65:7084-7084(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31964; AAA46152.1; -
DR PIR; A36770; CGBEHS.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF01391; Collagen; 1.
KW Collagen; Repeat.
FT DOMAIN 15 70
FT SEQUENCE 105 AA; 10260 MW; EF3DF0FE0FB446F0 CRC64;
SQ
Query Match 9.2%; Score 98; DB 1; Length 105;
Best Local Similarity 37.2%; Pred. No. 0.58;
Matches 29; Conservative 5; Mismatches 30; Indels 14; Gaps 4;
QY 1 MAKOPS--DVSSECDREGLOLQPAERPPQLRPGAPTSLOTPEQG--MPEGNHGEGDSCP 56
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MASEPNRYRTEETGDRGPGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 52
QY 57 HG--SPQGLAPPASP 72
Db ||||||:|||||
53 QGPPGPGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 70
RESULT 12
CUTL1_HUMAN
ID CUTL1_HUMAN STANDARD; PRT; 1505 AA.
AC P39880; O9UEV5;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE "Human CCAAT displacement protein (CDP) (Cut-like 1).
GN CUTL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Umbilical vein;
RX MEDLINE=93250973; PubMed=1301999;
RA Neufeld E.J., Skalniak D.G., Lievens P.M.-J., Orkin S.H.;
RT "Human CCAAT displacement protein is homologous to the Drosophila
RT homeoprotein, cut.";
RL Nat. Genet. 1:50-55(1992).
RN [2]
RP SEQUENCE OF 48-224 FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
CC -!- FUNCTION: PROBABLY HAS A BROAD ROLE IN MAMMALIAN DEVELOPMENT AS A
CC REPRESSOR OF DEVELOPMENTALLY REGULATED GENE EXPRESSION. MAY ACT BY
CC PREVENTING BINDING OF POSITIVELY-ACTIVATING CCAAT FACTORS TO
CC PROMOTERS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 (SHOWN HERE) AND 2;

```

```

CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- DOMAIN: ASN AT POSITION 47 OF THE HOMEBOX MAY PARTICIPATE IN
CC REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
CC HETERODIMERIZATION.
CC -!- SIMILARITY: CONTAINS 3 CUT DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CUT FAMILY OF HOMEBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M74099; -: NOT_ANNOTATED_CDS.
DR EMBL; AF047825; AAC78778.1; -
DR HSSP; P10037; IAU7.
DR TRANSFAC; T00100; -.
DR MIM; 116896; -.
DR InterPro; IPR003350; CUT.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF02376; CUT; 3.
DR Pfam; PF00046; homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW Transcription regulation; Homeobox; DNA-binding;
KW Developmental protein; Nuclear protein; Repeat; Repressor;
KW Coiled coil; Alternative splicing.
FT DOMAIN 7 363
FT DNA_BIND 542 629
FT DOMAIN 669 725
FT DNA_BIND 934 1021
FT DNA_BIND 1117 1204
FT DNA_BIND 1244 1303
FT VARSPLIC 632 653
FT SEQUENCE 1505 AA; 164353 MW; 860E14D508D4DE11 CRC64;
SQ
Query Match 9.2%; Score 97.5; DB 1; Length 1505;
Best Local Similarity 23.2%; Pred. No. 8.6;
Matches 52; Conservative 25; Mismatches 82; Indels 65; Gaps 10;
QY 6 SDVSECDREGLOLQPAERPPQLRPGAPTSLOTPEQGNPEGNHGEGDSCP 61
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
404 SLDLGSARRKGD-QPESRRPGSLPAPPSPQPNPGEASNTNGTHQFSPAGLSQDFFS 462
QY 62 ----GPIAPPASPGPFATRSPLFIFMRR--SSLLSR-----SSGYFSFD--TD 102
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 SSLASPSLPLASTGKFALNS---LLQRQLMSQFSYKAMQAEAGTSMIFSTGYSTNSS 519
QY 103 RSPAPMSCDKSTQTPSP-----COAFNHVLS 129
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 QSPLOQSPDVNGMAPSPSQSESAGSVSEGEEMDTAETARQVKQOLIKHNIGQRIFGHTVL 579
QY 130 AMASMRQAEADMRPEIWIQAELRRIGDEFNAYYARRVFLNNYQ 173
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
580 GLSGSGVSEIL-ARPKPNKLTVR--GKE--PFHKMKQFLSDEQ 618
RESULT 13
DLG2_RAT
ID DLG2_RAT STANDARD; PRT; 852 AA.
AC Q63622; Q62939; P70548;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Channel associated protein of synapse-110 (Chapsyn-110) (Synaptic
DE density protein PSD-93) (Discs, large homolog 2).
GN DLG2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96310881; PubMed=8755482;
RA Kim E., Cho K.O., Rothschild A., Sheng M.;
RT "Heteromultimerization and NMDA receptor-clustering activity of
Chapsyn-110, a member of the PSD-95 family of proteins.";
RL Neuron 17:103-113(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96193770; PubMed=8625413;
RA Brenman J.E., Chao D.S., Gee S.H., McGee A.W., Craven S.E.,
RA Santillano D.R., Wu Z., Huang F., Xia H., Peters M.F.,
RA Froehner S.C., Bredt D.S.;
RT "Interaction of nitric oxide synthase with the postsynaptic density
protein PSD-95 and alpha1-syntrophin mediated by PDZ domains.";
RL Cell 84:757-767(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Irie M., Hata Y., Takai Y.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNITS AS WELL AS POTASSIUM CHANNELS.
CC -!- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS A GUANYLATE KINASE-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U49049; AAB53243.1; .
DR EMBL; U50717; AAC52643.1; .
DR EMBL; U53368; AAB48562.1; .
DR HSSP; Q12959; 1PDR.
DR InterPro; IPR000619; Guanylate_kin.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00625; Guanylate_kin; 1.
DR Pfam; PF00595; PDZ; 3.
DR Pfam; PF00018; SH3; 1.
DR SMART; SM00072; Gukc; 1.
DR SMART; SM00228; PDZ; 3.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS0052; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS50106; PDZ; 3.
DR PROSITE; PS50002; SH3; 1.
DR SH3 domain; Repeat.
KW SH3 domain; Repeat.
FT DOMAIN 98 184 PDZ 1.
FT DOMAIN 193 279 PDZ 2.
FT DOMAIN 421 501 PDZ 3.
FT DOMAIN 536 606 SH3.
FT DOMAIN 662 852 GUANYLATE KINASE.
FT CONFLICT 181 182 VR -> IL (IN REF. 2).
FT CONFLICT 228 228 I -> M (IN REF. 2).
FT CONFLICT 326 326 R -> K (IN REF. 2).
FT CONFLICT 339 339 D -> E (IN REF. 3).
FT CONFLICT 450 454 MISSING (IN REF. 2).
FT CONFLICT 464 465 GD -> RK (IN REF. 2).
FT CONFLICT 474 474 D -> H (IN REF. 2).
FT CONFLICT 476 476 R -> P (IN REF. 2).
FT CONFLICT 478 478 A -> D (IN REF. 2).
FT CONFLICT 484 486 AAA -> LP (IN REF. 2).
FT CONFLICT 506 506 A -> S (IN REF. 2).
FT CONFLICT 569 569 H -> N (IN REF. 2).
FT CONFLICT 586 586 L -> Q (IN REF. 2).

FT CONFLICT 626 641 MISSING (IN REF. 2).
FT CONFLICT 639 639 K -> A (IN REF. 3).
FT CONFLICT 726 726 F -> L (IN REF. 1).
FT CONFLICT 733 733 N -> Y (IN REF. 2).
FT CONFLICT 749 749 E -> V (IN REF. 1).
FT CONFLICT 756 756 L -> H (IN REF. 2).
FT CONFLICT 791 792 KR -> NG (IN REF. 2).
FT CONFLICT 794 794 T -> M (IN REF. 2).
SQ SEQUENCE 852 AA; 94934 MW; F8D414A8B9CF5B09 CRC64;

Query Match 9.1%; Score 97; DB 1; Length 852;
Best Local Similarity 23.8%; Pred. No. 5.3;
Matches 44; Conservative 21; Mismatches 58; Indels 62; Gaps 9;

QY 28 LRPGAPTSI-QTEPCQNPENHGEGDSCPH--GSPQGPI-----APPASPGPFA----- 74
Db LVKGRPTTIYMDPYGPPDITHSYPPMENHLLSGNNGTLEYKTSLPPIISPGRYSPKX 333

QY 75 -----TRSPFIEMRRSSLLSRSSGYFSFDTDRSPAPMSCDKS--TQTPSPPCQAFN 125
Db 334 MLVEDDYTRPPEPVYSTVKNKCDKPAS-----PRHYSPEVCDKSFLLSTPY----- 380

QY 126 HVL-----SAMASMRQAEPADMRPEIWIQAELRRIGDEFNAYYARVFNLYNQAAEDHP 179
Db 381 HVHGLLPDSDMTSHSOHSTATRPSVTLQRAI-----SLEGE 419

QY 180 RVVIL 184
Db 420 RKVVL 424

RESULT 14
DLG2_HUMAN STANDARD; PRT; 870 AA.
AC Q15700;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Channel associated protein of synapse-110 (Chapsyn-110) (Discs, large
DE homolog 2).
GN DLG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96310881; PubMed=8755482;
RA Kim E., Cho K.O., Rothschild A., Sheng M.;
RT "Heteromultimerization and NMDA receptor-clustering activity of
Chapsyn-110, a member of the PSD-95 family of proteins.";
RL Neuron 17:103-113(1996).
CC -!- FUNCTION: INTERACTS WITH THE CYTOPLASMIC TAIL OF NMDA RECEPTOR
CC SUBUNITS AS WELL AS POTASSIUM CHANNELS.
CC -!- SIMILARITY: CONTAINS 3 PDZ/DHR DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE MAGUK FAMILY OF CELL JUNCTION PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32376; AAB04949.1; .
DR HSSP; Q12959; 1PDR.
DR MIM; 603583; .
DR InterPro; IPR000619; Guanylate_kin.
DR InterPro; IPR001478; PDZ.


```
DR EMBL; L22545; AA19787.1; -.
DR PDB; 1KOE; 16-FEB-99.
DR MGD; MGI:88451; Col18a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF02210; TSPN; 2.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
KW 3D-structure.
FT SIGNAL          1   26
FT CHAIN           27 1527
FT CHAIN           1344 1527
FT CHAIN           27 538
FT DOMAIN          539 585
FT DOMAIN          566 575
FT DOMAIN          576 649
FT DOMAIN          650 673
FT DOMAIN          674 795
FT DOMAIN          796 818
FT DOMAIN          819 901
FT DOMAIN          902 915
FT DOMAIN          916 957
FT DOMAIN          958 970
FT DOMAIN          971 1043
FT DOMAIN          1044 1053
FT DOMAIN          1084 1086
FT DOMAIN          1087 1098
FT DOMAIN          1099 1122
FT DOMAIN          1123 1129
FT DOMAIN          1130 1181
FT DOMAIN          1182 1194
FT DOMAIN          1195 1212
FT DOMAIN          1213 1527
FT CARBOHYD       338 338
FT CARBOHYD       700 700
FT DISULFID       1376 1516
FT DISULFID       1478 1508
FT SITE           1104 1106
FT VARSPLIC       1 212
FT VARSPLIC       213 238
FT CONFLICT       900 900
FT CONFLICT       947 947
FT CONFLICT       964 964
FT CONFLICT       1157 1157
FT CONFLICT       1266 1266
FT CONFLICT       1276 1276
FT CONFLICT       1437 1437
SQ SEQUENCE        1527 AA; 156008 MW; 9645045AF140B513 CRC64;

POTENTIAL.
COLLAGEN ALPHA 1(XVIII) CHAIN.
ENDOSTATIN.
NONHELICAL REGION 1 (NC1).
TRIPLE-HELICAL REGION 1 (COL1).
NONHELICAL REGION 2 (NC2).
TRIPLE-HELICAL REGION 2 (COL2).
NONHELICAL REGION 3 (NC3).
TRIPLE-HELICAL REGION 3 (COL3).
NONHELICAL REGION 4 (NC4).
TRIPLE-HELICAL REGION 4 (COL4).
NONHELICAL REGION 5 (NC5).
TRIPLE-HELICAL REGION 5 (COL5).
NONHELICAL REGION 6 (NC6).
TRIPLE-HELICAL REGION 6 (COL6).
NONHELICAL REGION 7 (NC7).
TRIPLE-HELICAL REGION 7 (COL7).
NONHELICAL REGION 8 (NC8).
TRIPLE-HELICAL REGION 8 (COL8).
NONHELICAL REGION 9 (NC9).
TRIPLE-HELICAL REGION 9 (COL9).
NONHELICAL REGION 10 (NC10).
TRIPLE-HELICAL REGION 10 (COL10).
NONHELICAL REGION 11 (NC11).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).

CELL ATTACHMENT SITE (POTENTIAL).
MISSING (IN SHORT ISOFORM).
AVPTQLPPFQSNLQAPLGRPSAPPDF -> MAPRWHLLDVL
TSLVLLLVARVSNAE (IN SHORT ISOFORM).
P -> L (IN REF. 4).
P -> F (IN REF. 4).
A -> R (IN REF. 4).
R -> P (IN REF. 4).
P -> L (IN REF. 4).
L -> F (IN REF. 4).
L -> V (IN REF. 4).

Query Match          9.0%; Score 96; DB 1; Length 1527;
Best Local Similarity 31.9%; Pred. No. 11;
Matches 36; Conservative 5; Mismatches 40; Indels 32; Gaps 6;

QY 16 GROLQPAERPPQLRPGAPTSLQTEPGNPEGNHGGSDSCPFG-----SPQGPLAPPA 68
   || : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 594 GRDGEFGDGEDGRPG-----DTGPGQGF---GTPGDVKGKGDGPGIGRPGPPPG 644
   || : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 69 SPGPFATRSPL-FIFMRRSLLSRSSSGYFSDTDRSPAPMSCDKSTQTPSP 120
   ||| : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 645 PPGPSFRQDKLTFIDMEGSG-----FSGDIESLRGRG-----FPGPP 682
```

Search completed: August 16, 2002, 10:05:14
Job time: 232 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 16, 2002, 10:01:02 ; Search time 26.66 Seconds
(without alignments)
1284.809 Million cell updates/sec

Title: US-09-508-832-10
Perfect score: 1065
Sequence: 1 MAKOPSDVSSECDREGRLQ.....PRMVLRLRLVIVRLVWRMH 198

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_protent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	110.5	10.4	472	12	O36424	O36424 alcelaphine
2	106.5	10.0	620	12	P89435	P89435 herpes simp
3	106.5	10.0	783	4	O9NXXD0	O9NXXD0 homo sapien
4	106.5	10.0	950	5	O9VE45	O9VE45 drosophila
5	105.5	9.9	574	3	O36027	O36027 schizosacch
6	103.5	9.7	556	2	O92FS8	O92FS8 streptomyc
7	103.5	9.7	556	2	O9S2C0	O9S2C0 streptomyc
8	103.5	9.7	1051	4	O9S135	O9S135 homo sapien
9	102.5	9.6	1517	11	O91ZD2	O91ZD2 mus musculus
10	102	9.6	321	4	O96SQ7	O96SQ7 homo sapien
11	101	9.5	1820	13	O91907	O91907 pagrus ma'jo
12	100.5	9.4	620	12	O89893	O89893 human herpe
13	99.5	9.3	280	5	O9NFZ9	O9NFZ9 ascaris suu
14	99.5	9.3	378	11	O9CSY5	O9CSY5 mus musculus
15	99.5	9.3	874	3	O96VJ2	O96VJ2 pneumocysti
16	99	9.3	257	4	O9UDU7	O9UDU7 homo sapien

Q96ha7 homo sapien
Q94gk7 oryza sativ
O62432 caenorhabdi
Q60735 mus musculu
Q95043 homo sapien
Q93174 arabidopsis
Q93107 acanthamoeb
Q9qx49 mus musculu
Q94dl2 oryza sativ
Q61434 mus musculu
Q62001 mus musculu
Q00487 hydra sp. m
Q97737 bos taurus
Q9f691 streptococc
Q99m33 mus musculu
Q60749 mus musculu
O88539 mus musculu
Q923d5 mus musculu
Q9jka7 rattus norv
Q98tw1 brachydanio
O9erc1 rattus norv
O9jjs4 rattus norv
O22514 santalum al
Q9vpp1 drosophila
Q96c67 homo sapien
Q99495 homo sapien
Q99621 homo sapien
Q84187 measles vir
Q99in4 canis famil

ALIGNMENTS

RESULT 1
O36424 PRELIMINARY; PRT; 472 AA.
ID O36424;
AC O36424;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE PUTATIVE GLYCOPROTEIN.
OS Alcelaphine herpesvirus 1 (Wildebesset herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gamaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=35252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C500;
RX MEDLINE=97404659; PubMed-9261371;
RA Ensser A., Pflanz R., Fleckenstein B.;
RT "Primary structure of the alcelaphine herpesvirus 1 genome.";
RL J. Virol. 71:6517-6525(1997).
DR EMBL; AF005370; AAC58121.1; -.
SQ SEQUENCE 472 AA; 50142 MW; 0DDEF854764B8342 CRC64;

Query Match 10.4%; Score 110.5; DB 12; Length 472;
Best Local Similarity 22.9%; Pred. No. 0.0067;
Matches 48; Conservative 19; Mismatches 78; Indels 65; Gaps 7;
QY 25 PPOLRPGAPTSL-----QTEPGNPEGNH----- 48
Db 279 PGCTKPCPTPLWDPCAVLLSSGSTPPWDRRKILHRTKVTGTPTVTPVYLISDDDL 338
QY 49 ---GGECDSCPHGSPQGLAPPASPGFATRPLIFPMRRSLLSRSSSGYFSDTDRSP 105
Db 339 DRKGAAG-----GGPKKPLPSRDPAGSGQGRGP-----TPQSKESTSSGLSGATPRSK 386
QY 106 APMSCKDKSTQTPSPPCQAFNHLYLSAMASMRQAPADMRPEIW-IAQLRIRIGDFENYAYA 164
Db 387 APAGRGRQEPPLPP-----RDYDPASPEALRLRLDITVQKRVYIGGAYNELEA 435

RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpén G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003723; AAF5584.1; -.
DR FlyBase: FBgn038642; CG7709.
DR InterPro: IPR002965; P_rich_extensin.
DR InterPro: IPR004019; YLP.
DR Pfam: PF02757; YLP; 3.
DR PRINTS: PR01217; PRICEXTENSIN.
SQ SEQUENCE 950 AA; 91280 MW; B57AB0EA4D30D3D7 CRC64;

Query Match 10.0%; Score 106.5; DB 5; Length 950;
Best Local Similarity 30.3%; Pred. No. 0.038;
Matches 37; Conservative 14; Mismatches 58; Indels 13; Gaps 4;
QY 30 PGAPTSLOTEPGNPNHGEGDSCPHGSPQGLAPPASP--GPFATSPFLIFMRSS 87
Db 785 PRAPSSSYAAP--SFSSSGGYPYAPAPSNSYAGTAPSSSYGAPSGSPFSAPSSSY 842
QY 88 LLRSRSSGYFSFDTDRSP-APMSCDKSTQTPPPCQAF-----NHLYSAMASMRQAE 138
Db 843 GAPSTGCGSSGSSSSSSSFGASSSSSAGYPSAPSSSYGAPSTGSGHSFSSAPSSSYAP 902
QY 139 PA 140
Db 903 PA 904
RESULT 5
ID O36027 PRELIMINARY; PRT; 574 AA.
AC O36027;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE WISKOTT-ALDRICH SYNDROME PROTEIN HOMOLOG 1.

GN WSP1 OR SPAC4F10.15C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=JS21;
RA Zankel T.C., Ow D.W.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST LAS17.
DR EMBL: AF038575; AAB92587.1; -.
DR EMBL: Z98980; CAB11718.1; -.
DR InterPro: IPR000697; RanBP1_WASP.
DR InterPro: IPR001960; WH1.
DR InterPro: IPR003124; WH2.
DR Pfam: PF00568; WH1; 1.
DR Pfam: PF02205; WH2; 1.
DR SMART: SM00461; WH1; 1.
FT DOMAIN 311 317 POLY-PRO.
FT DOMAIN 337 343 POLY-PRO.
FT DOMAIN 361 366 POLY-PRO.
FT VARIANT 248 248 L -> V (IN STRAIN JS21).
SQ SEQUENCE 574 AA; 59605 MW; C6E5EFC6A02F0E4 CRC64;
Query Match 9.9%; Score 105.5; DB 3; Length 574;
Best Local Similarity 26.8%; Pred. No. 0.027;
Matches 41; Conservative 17; Mismatches 66; Indels 29; Gaps 6;
QY 21 PAERPPQLRPGAPTSLOTEPGNPEG-----NHGEGDSCPHGSPQGLAPPASPGP--- 72
Db 424 PSAPPSLPSPAPPSL---PMGAPAAPPLPPSAPIAPPLPAGMPAAPPLPAAAPAPPAP 480
QY 73 -FATRSPLIFM-----RRSLLSRSSSGYFSFDTDRSPAPMSCDKSTQTPSPCQA- 123
Db 481 APAPAAPVASIAELPOODGRANLMASIRASGGMDLKRKVSASPSVASTKTSNPPEAP 540
QY 124 -FNHLYSAMAS-----MRQAEPADMRPEIW 147
Db 541 PSNNLMDALASALNQRTKVAQSDDEDDDEW 573
RESULT 6
ID Q9ZFS8 PRELIMINARY; PRT; 556 AA.
AC Q9ZFS8;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SERINE/THREONINE PROTEIN KINASE.
GN PKSC.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RC SEQUENCE FROM N.A.
RA Bakal C.J., Davies J.E.;
RT "Cloning, nucleotide sequence and expression of a serine/threonine
protein kinase gene from Streptomyces coelicolor.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF094711; AAC64406.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.

[illegible]

RA Bertram P., Jochum W., Barrandon Y., Busslinger M.;
RT "The transcriptional repressor CDP (Cutli) is essential for epithelial
cell differentiation of the lung and the hair follicle.";
RL Genes Dev. 15:2307-2319(2001).
DR EMBL: AY037807; AAKS9986.1; -;
SQ SEQUENCE 1517 AA; 163852 MW; B17A8740621EBBAC CRC64;

Query Match 9.6%; Score 102.5; DB 11; Length 1517;
Best Local Similarity 23.7%; Pred. No. 0.17; Mismatches 25; Conservative 54; Gaps 9;

QY 6 SDVSSCDREGQLQPAERPPQLRPGAPTSLOT---EPQNPENHGEGDSCPHGSPQG 62

Db 404 SDSLGSARRKGRDQPSRRPGLPASPPQLPNTGCEQVSNTNGTH-----HFPAG 455

QY 63 -----PLAPPASGPFATRPSLFIFMRRS-----SLLSRSSGYSFSD 100

Db 456 LSQDFSSNLASPLSTAGFALNSLLQRLQMLQSFYSKAMQEAAGSTSTFSTGPGYSTN 515

QY 101 TDRSPAM--SCDKSTQTPSP-----COAFN 125

Db 516 STSSPSPLQSDPDVNGMSPQSESAGSISBGEIDTAEIARQVKEQLIKHNIGQRIFG 575

QY 126 HYLASAMASMRQAEADMRPEIWIQAELRIGDEFNAYARRVFLNNYQ 173

Db 576 HVLGLSGSVSEIL-ARPKWNKLTVR--GKE--PFKMKQFLSDEQ 618

RESULT 10

ID Q96SQ7 PRELIMINARY; PRT; 321 AA.
AC Q96SQ7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE CDNA FLJ14708 FIS, CLONE NT2RP300603, WEAKLY SIMILAR TO
DE NEUROGENIC DIFFERENTIATION FACTOR 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuho Y., Kanehori K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK027614; BAB5233.1; -;
SQ SEQUENCE 321 AA; 34627 MW; 8FB9C04C23BA4258 CRC64;

Query Match 9.6%; Score 102; DB 4; Length 321;
Best Local Similarity 23.8%; Pred. No. 0.03; Mismatches 21; Conservative 53; Gaps 9;

QY 2 AKQPSDVSSECDREGQLQPAERPPQLRPGAPTSLOTPEQNPENHGEGDSCPHGSPQ 61

Db 101 APEVSDARKRCFALGAVGPGLETPPPPPPPAP---QSQAPGGPEAQ-----PPREP- 148

QY 62 GP-----LAPASGPFATRPSLFIFMRRSLLSRSSGYSFSDRSPAPMSCDKSTQ 115

Db 149 GPRPILLCAPPARPAPSA-----PPAPPAPPSTV 179

QY 116 TSPSPCCQ-----AFNHYLSAMASMRQ--AEPADMRPEIWIQAELRRIGDEFN- 160

Db 180 RPAPTRPGESSYSSISHVIYNNHQDSSASPRKRPFCEATAASSEIKALQOTRLLANARE 239

QY 161 -----AYARRVFLNNYQAAEDHPRVILRL-LRYIVRL 193

Db 240 RTRVHTISAFAEALRKQVPCYSYGQKLSKLAILRIACNVLSL 282

RESULT 11

ID Q91907 PRELIMINARY; PRT; 1820 AA.
AC Q91907;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PRO-ALPHA 1 TYPE V/XI COLLAGEN.
GN COLV/XIAl.
OS Pagrus major (Red sea bream) (Chrysophrys major).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Pagrus;
OX NCBI_TaxID=143350;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21240220; PubMed=11342118;
RA Touhata K., Tanaka H., Yokoyama Y., Sakauechi M., Toyohara H.;
RT "Structure of a full-length cDNA clone for the pro-1(V/XI) collagen
chain of red seabream";
RL Biochim. Biophys. Acta 1517:323-326(2001).
DR EMBL: AB045975; BAB03287.1; -;
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000885; Fib_collagen_C.
DR InterPro: IPR001791; Laminin_G.
DR InterPro: IPR001230; Prenyltn.
DR InterPro: IPR003129; TSPN.
DR Pfam: PF01410; COLFI; 1.
DR Pfam: PF02210; TSPN; 1.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
DR SMART: SM00282; Lamg; 1.
DR SMART: SM00210; TSPN; 1.
DR PROSITE: PS00294; PRENYLATION; UNKNOWN_1.
KW Collagen.
SQ SEQUENCE 1820 AA; 181678 MW; 46E45E8AF7AD3DAE CRC64;

Query Match 9.5%; Score 101; DB 13; Length 1820;
Best Local Similarity 29.4%; Pred. No. 0.29; Mismatches 50; Conservative 8; Indels 54; Gaps 9;

QY 30 PGAPTSLOTPEQNPENHGEGDSCPHGSPQGLAPPASGPFATRPSLFIFMRRSLL 89

Db 1474 PGEPGEKGRGLPGPQGTGGKDGSTVG--PGPLGPPGPPG-----ISGPQCKG 1523

QY 90 SRSSSGYFSFDTRD---PAPMSCDKSTQTPSP---COAFNHYLSAMASMRQAEADM 142

Db 1524 SKGSTGTQCKGNDNGIIGPPG-----PGPGDIIPLPMQQSARKTRQAEQ-- 1572

QY 143 RPEIWIQAELRRIGDEFNAYYA-----RRVF--LNNYQAAEDHPRM 181

Db 1573 -----GDEAMADYGVGVGMEGMEDVFGSLNNLK--QDIERM 1607

RESULT 12

ID Q89893 PRELIMINARY; PRT; 620 AA.
AC Q89893;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE HYPOTHETICAL 67.1 KDA PROTEIN.
GN DR2.
OS Human herpesvirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10368;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=95266321; PubMed=7747482;
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
Martin M.E., Efsthathiou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
and genome evolution.";
RL Virology 209:29-51(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=90080133; PubMed=2152817;
RA Lawrence G.L., Chee M., Craxton M.A., Gompels U.A., Honess R.W.,
Barrell B.G.;
RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";
RL J. Virol. 64:287-299(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=91237802; PubMed=1851860;
RA Chang C.K., Balachandran N.;
RT "Identification, characterization, and sequence analysis of a cDNA
encoding a phosphoprotein of human herpesvirus 6.";
RL J. Virol. 65:2884-2894(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=91333007; PubMed=1651403;
RA Teo I.A., Griffin B.E., Jones M.D.;
RT "Characterization of the DNA polymerase gene of human herpesvirus 6.";
RL J. Virol. 65:4670-4680(1991).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=91226542; PubMed=1851252;
RA Thomson B.J., Efsthathiou S., Honess R.W.;
RT "Acquisition of the human adeno-associated virus type-2 rep gene by
human herpesvirus type-6.";
RL Nature 351:78-80(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=91374590; PubMed=1654446;
RA Martin M.E., Nicholas J., Thomson B.J., Newman C., Honess R.W.;
RT "Identification of a transactivating function mapping to the putative
immediate-early locus of human herpesvirus 6.";
RL J. Virol. 65:5381-5390(1991).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=92333249; PubMed=1321206;
RA Efsthathiou S., Lawrence G.L., Brown C.M., Barrell B.G.;
RT "Identification of homologues to the human cytomegalovirus US22 gene
family in human herpesvirus 6.";
RL J. Gen. Virol. 73:1661-1671(1992).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=92148942; PubMed=1310766;
RA Geng Y., Chandran B., Josephs S.F., Wood C.;
RT "Identification and characterization of a human herpesvirus 6 gene
segment that trans activates the human immunodeficiency virus type 1
promoter.";
RL J. Virol. 66:1564-1570(1992).
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=93091236; PubMed=1333836;
RA Gompels U.A., Carss A.L., Sun N., Arrand J.R.;
RT "Infectivity determinants encoded in a conserved gene block of human
herpesvirus-6.";
RL DNA Seq. 3:25-39(1992).
RN [10]

RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=92260671; PubMed=1374813;
RA Neipel F., Ellinger K., Fleckenstein B.;
RT "Gene for the major antigenic structural protein (p100) of human
herpesvirus 6.";
RL J. Virol. 66:3918-3924(1992).
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=92333248; PubMed=1321205;
RA Thomson B.J., Honess R.W.;
RT "The right end of the unique region of the genome of human herpesvirus
6 U1102 contains a candidate immediate early gene enhancer and a
homologue of the human cytomegalovirus US22 gene family.";
RL J. Gen. Virol. 73:1649-1660(1992).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=93187613; PubMed=8383182;
RA Ellinger K., Neipel F., Foa-Tomasi L., Campadelli-Fiume G.,
Fleckenstein B.;
RT "The glycoprotein B homologue of human herpesvirus 6.";
RL J. Gen. Virol. 74:495-500(1993).
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=93224882; PubMed=8385692;
RA Gompels U.A., Carrigan D.R., Carss A.L., Arno J.;
RT "Two groups of human herpesvirus 6 identified by sequence analyses of
laboratory strains and variants from Hodgkin's lymphoma and bone
marrow transplant patients.";
RL J. Gen. Virol. 74:613-622(1993).
RN [14]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=93389439; PubMed=8397282;
RA Liu D.X., Gompels U.A., Nicholas J., Lelliott C.;
RT "Identification and expression of the human herpesvirus 6 glycoprotein
H and interaction with an accessory 40K glycoprotein.";
RL J. Gen. Virol. 74:1847-1857(1993).
RN [15]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=94025558; PubMed=7692666;
RA Liu D.X., Gompels U.A., Foa-Tomasi L., Campadelli-Fiume G.;
RT "Human herpesvirus-6 glycoprotein H and L homologs are components of
the gp100 complex and the gH external domain is the target for
neutralizing monoclonal antibodies.";
RL Virology 197:12-22(1993).
RN [16]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=93331710; PubMed=7687803;
RA Pellett P., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E.,
Greenamoyer C., Dambaugh T.R.;
RT "A strongly immunoreactive virion protein of human herpesvirus 6
variant B strain 429: identification and characterization of the gene
and mapping of a variant-specific monoclonal antibody reactive
epitope.";
RL Virology 195:521-531(1993).
RN [17]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102;
RX MEDLINE=93323202; PubMed=7687301;
RA Pfeiffer B., Berneman Z.N., Neipel F., Chang C.K., Tirwatnapong S.,
Chandran B.;
RT "Identification and mapping of the gene encoding the glycoprotein
complex gp82-gp105 of human herpesvirus 6 and mapping of the
neutralizing epitope recognized by monoclonal antibodies.";
RL J. Virol. 67:4611-4620(1993).
RN [18]
RP SEQUENCE FROM N.A.

RC STRAIN-U1102;
RX MEDLINE=95146989; PubMed=7844567;
RA Gompels U.A., Macaulay H.A.;
RT "Characterization of human telomeric repeat sequences from human
herpesvirus 6 and relationship to replication.";
RL J. Gen. Virol. 76:451-458(1995).
RN [19]
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE=94047392; PubMed=8230490;
RA Dewhurst S., Dollard S.C., Pellett P.E., Dambaugh T.R.;
RT "Identification of a lytic-phase origin of DNA replication in human
herpesvirus 6B strain 229.";
RL J. Virol. 67:7680-7683(1993).
RN [20]
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE=94118404; PubMed=8289364;
RA Nicholas J., Martin M.;
RT "Nucleotide sequence analysis of a 38.5-kilobase-pair region of the
genome of human herpesvirus 6 encoding human cytomegalovirus
immediate-early gene homologs and transactivating functions.";
RL J. Virol. 68:597-610(1994).
RN [22]
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE=94202284; PubMed=8151768;
RA Schiwe U., Neipel F., Schreiner D., Fleckenstein B.;
RT "Structure and transcription of an immediate-early region in the human
herpesvirus 6 genome.";
RL J. Virol. 68:2978-2985(1994).
RN [23]
RP SEQUENCE FROM N.A.
RC STRAIN-U1102;
RX MEDLINE=94181269; PubMed=8134119;
RA Thompson J., Choudhury S., Kashanchi F., Doniger J., Berneman Z.,
RA Frenkel N., Rosenthal L.J.;
Query Match 9.4%; Score 100.5; DB 12; Length 620;
Best Local Similarity 31.0%; Pred. No. 0.093;
Matches 36; Conservative 10; Mismatches 51; Indels 19; Gaps 5;
QY 21 PAERPPQLRP-----GAPTSIQTEPQGNHNG-----GGGDCSPHGSPGGLA 65
DB 486 PAEPPTCTPRPPRPSSTPLSAVSRPSAPPVPPPPSTARVRFFLSSSSSSPSYSP-APLS 544
QY 66 PPASPGPEATRSLFFIMRRSSLLS--RSSSGY-FSFDTRSPAPMSCDKSTQTPS 118
DB 545 PPSVPSSPSRSPFPIPRKSLGLRAKPRVSSGHPAAPPAPSPAPARSERVTSVPS 600
RESULT 13
Q9NFZ9 PRELIMINARY; PRT; 280 AA.
AC Q9NFZ9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PUTATIVE CUTICULAR COLLAGEN (FRAGMENT).
GN COL-D.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RT Di Mito C., Betschart B.;
RT "Identification of two collagen genes from the cuticle of Ascaris

RT suum.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AT249373; CAB85466.1; -;
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR002486; Col_cuticle_N.
DR Pfam: PF01391; Collagen; 2.
DR Pfam: PF01484; Col_cuticle_N; 1.
KW Collagen.
FT NON_TER 1
SQ SEQUENCE 280 AA; 27796 MW; E61086529408BC10 CRC64;
Query Match 9.3%; Score 99.5; DB 5; Length 280;
Best Local Similarity 30.4%; Pred. No. 0.046;
Matches 34; Conservative 9; Mismatches 36; Indels 33; Gaps 6;
QY 15 EGRQLQPAERPPQLRPGAPTSLQTEPQGNP--EGNIGGEGDCSPHGSP--QGFLAPASP 70
DB 193 EAQQLIPGERPPPGPPGPG--RPGPPGFRGPGKDGNDGNGPPGPGRPPGQGN 250
QY 71 GPFATRSPLFIEMRRSSLLSSSSGYFSFDTDRSPAPMSCDKSTQTPSPPCQ 122
DB 251 G-----QQGRPG-----EPGQAGAPGSCDH-----CPCPAR 275
RESULT 14
Q9CSY5 PRELIMINARY; PRT; 378 AA.
AC Q9CSY5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 2610034E13RIK PROTEIN (FRAGMENT).
GN 2610034E13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Carninci P., de Bonaldo M.F.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK011662; BAB27764.1; -;
DR MGD; MGI:1922863; 2610034E13RIK.
FT NON_TER 1
SQ SEQUENCE 378 AA; 40207 MW; DDD783BA958F104C CRC64;
Query Match 9.3%; Score 99.5; DB 11; Length 378;
Best Local Similarity 31.8%; Pred. No. 0.066;
Matches 42; Conservative 10; Mismatches 31; Indels 49; Gaps 8;
QY 16 GROLQPAERPPQ--LRPGAPTSLQTEPQG-----NPEGNHGEGDCSPHGSPQG 62

Search completed: August 16, 2002, 10:04:56
Job time: 234 sec